

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2002, 00:20:11 ; Search time 3375.48 Seconds
(without alignments)
11345.220 Million cell updates/sec

Title: US-09-755-456-1

Perfect score: 1830
Sequence: 1 gaattcagatgcctcatcac.....tcccttagaagtgtggtac 1830

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_stg: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_cm: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_stg: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_inv: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length DB	ID	Description
------------	-------------	-----------	----	-------------

1	1830	100.0	1830	6	AX007049	AX007049 Sequence
2	1830	100.0	1831	3	ECU5666	AJ005666 Encephal1
3	1828.4	99.9	220294	3	CNS076ED	AL590446 Chromosom
4	1019	55.7	1019	3	AF310677	AF310677 Encephal1
5	1011.6	55.3	1018	3	AF310678	AF310678 Encephal1
6	844.8	46.2	936	3	AF310679	AF310679 Encephal1
7	176	9.6	1116	6	AX007051	AX007051 Sequence
8	144.2	7.9	1696	3	AF044915	AF044915 Encephal1
9	102.6	5.6	1340	3	AF338363	AF338363 Encephal1
10	101.4	5.5	1184	3	AF338361	AF338361 Encephal1
11	99.4	5.4	1297	3	AF338362	AF338362 Encephal1
12	99	5.4	1232	3	AY024342	AY024342 Encephal1
13	49.4	2.7	125020	9	AF429315	AF429315 Homo sap1
14	47.4	2.6	125020	9	AF429315	AF429315 Homo sap1
15	46.6	2.5	203108	9	CNS05TFC3	AL355032 Human chr
16	46.6	2.5	207774	2	CNS0000H	AL049836 Human chr
17	46	2.5	66491	2	AC006715	AC006715 Caenorhab
18	46	2.5	89370	3	AC084153	AC084153 Caenorhab
19	44.4	2.4	686	8	CNS01DIR	AL116868 Botrytis
20	43.2	2.4	33499	9	AL590548	AL590548 Human DNA
21	43.2	2.4	163057	2	AC012419	AC012419 Homo sap1
22	43.2	2.4	163852	9	AC015971	AC015971 Homo sap1
23	43	2.3	3489	9	AF217688S2	AF217689 Homo sap1
24	43	2.3	257817	2	AC006909	AC006909 Caenorhab
25	42.4	2.3	211579	2	AL645845	AL645845 Mus muscu
26	42.2	2.3	170707	2	AL589986	AL589986 Homo sap1
27	42.2	2.3	178671	2	AL627110	AL627110 Homo sap1
28	42.2	2.3	199739	2	AC005528	AC005528 Mus muscu
29	41.4	2.3	37369	3	CET26H2	282055 Caenorhabd1
30	41.4	2.3	56870	2	CET39B6-3	Continuation (4 of
31	41.4	2.3	110000	2	CET39B6-2	Continuation (3 of
32	40.6	2.2	7218	6	166494	166494 Sequence 14
33	40.4	2.2	34727	3	CET11A6	281498 Caenorhabd1
34	39.6	2.2	67453	2	AC104900	AC104900 Mus muscu
35	39.6	2.2	172224	2	AC087149	AC087149 Mus muscu
36	39.6	2.2	190738	2	AC087222	AC087222 Homo sap1
37	39.6	2.2	205228	2	AC068584	AC068584 Homo sap1
38	39.4	2.2	75153	2	AC100376	AC100376 Mus muscu
39	39.4	2.2	103829	9	AP000665	AP000665 Homo sap1
40	39.4	2.2	160222	9	AP002962	AP002962 Homo sap1
41	39.2	2.1	1204	3	AS0249373	AJ249373 Ascaris s
42	39.2	2.1	175231	2	AC097084	AC097084 Rattus no
43	39	2.1	696	8	CNS01A92	AL112638 Botrytis
44	39	2.1	9673	2	AC104348	AC104348 Homo sap1
45	39	2.1	143291	9	HS163G9	AL008733 Human DNA

ALIGNMENTS

RESULT 1	AX007049	1830 bp	DNA	linear	PAT 06-SEP-2000
LOCUS	AX007049	Sequence 1 from Patent WO0001724.			
DEFINITION	AX007049				
ACCESSION	AX007049.1	GI:994975			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Encephalitozoon cuniculi.				
REFERENCE	Encephalitozoon cuniculi				
AUTHORS	Eukaryota: Microsporidia; Unikaryonidae; Encephalitozoon.				
TITLE	1 (bases 1 to 1830)				
JOURNAL	Vivares, C., Danchin, A. and Delbac, F.				
FEATURES	Microsporidium polar tube proteins, nucleic acids coding for said				
source	Proteins and their uses				
BASE COUNT	460 a	486 c	434 g	450 t	
ORIGIN	/organism="Encephalitozoon cuniculi";				
	/db_xref="taxon:6035"				

Query Match 100.0%; Score 1830; DB 6; Length 1830;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaataagatgctcatcacttggagattaaaaattgatttcatattgtatatactcgt 60
DB 1 GAATTCAGATGCTCATACCTGGGATTTAAAAATTGATTCATTGTATATATACCTG 60
QY 61 ggggagagccggcgctcattcttcaagggtgtgcgcaaccagatgcacagaagtttc 120
DB 61 GGGGAGAGCCGGCTCGATTCATTCAGGGGtgcgcttacccttaccagtgacaggagtttc 120
QY 121 ggaaggtcttggatgtaagaagcatttgggttctcatccatgtcatcgtccct 180
DB 121 GGAGGtGCTTGGATGGAAGTAAAGCCATTGTGGTCTCATTCATGTCATGTCCTG 180
QY 181 ttgcgctgttcccaagaatccaattatctcccaaggacttcaacctcagaatgga 240
DB 181 TTCGGCTGTTTCCCAAGATCCAAATTTCTCCAGGACTTCACCCCTCAGAAATGGAAA 240
QY 241 cagaagatgaactctctgtgcaaatctgagatagatgagagacattgaaccacagag 300
DB 241 CAGAGATGAACCTCTGTGCAATGTAATTCATTCGATTTGGAGACATGGAACCCAGGAG 300
QY 301 ttgtaataaagaatataaataccctccgaacacgacagatttaagaatgaagttatc 360
DB 301 TTTGAATAAAGTATATAATACCTCCGAAACGACAGATTTAAGATGAAGATTTCTTA 360
QY 361 agatctctctgcctctatcttgcctgtatgaagtgtgaagattctatcaagaacgc 420
DB 361 AGATCTCTCTGCTCTATTCGCTTATGCGATGAAAGTGGAAATGCTATTCGCAACCGCAC 420
QY 421 tgtgcagcaatgcatatggccttaactccggagacagaggtatggttcagcagcgt 480
DB 421 TGTGCAGCAATGCAATATGGCTTAACCTCCGGACACAGGGTATGGCTCAGACCCGTCGT 480
QY 481 atgtgctgattcccaagacacccggagaaacataagcaaatctgtgcaacggttca 540
DB 481 ATGTGCTGATCCCGACACCCCGGGAACCATAGCAAACTGTGCAAGCGGTTTACAGAGACA 540
QY 541 catattctctcttcccgctgcacccacatctccagtgactccgggggaaaaactacg 600
DB 541 CATATTCTCTCTCCCGCTGCACCCACATCTCCAGTGCATCCGGGAAAACTAGGAGAGA 600
QY 601 atagaacatctccatcagctcgtcctgcagaagatgtagaacaatgcaagaattgcgt 660
DB 601 ATAGACATCTCCATGCGCTCCCTGCAGAAAGATAGAGAACATCAAGATTGCCGTATTGA 660
QY 661 agcactgcagacgacaggaacaacatacaggaacgacacagggcttctgtga 720
DB 661 AGCAGTCGCGACGACAGGAACAAATCAGGAGACACACAGGGGCTTTGTGAAA 720
QY 721 ccccaagagacagacagccttctgtcagtgatctccacacacactctgcgtacggt 780
DB 721 CCCCAAGACGACAGACCTTTGTGAGTATCTCCACCACTCCCGGTAACCGTGAATG 780
QY 781 tgaagctgcagagcttccatctgtgtgccaattgttccctgtcgttgcctcacccag 840
DB 781 TGAAGCTGCACAGCTCCATCTGTTGTGCGCACTGTCTCTGTCGTTCGCTCACACACAG 840
QY 841 cagttccagagctactaacaacatgaaacatccggtatctccgtgacagcaacagatc 900
DB 841 CAGTTCAGAGCTACTACAAATGAATGAAATCCGGTATTCTCTGACAGCAACAGATCTT 900
QY 901 ctggaactcttccccaagagacacttctgtgtcaggaacagagccatgactagctcgt 960
DB 901 CTGGCACTCTTCCCAAGGACCACTTGTGTGTCAGGACAGGCAATGCCATGCACTCTG 960
QY 961 gaacagaaacagatccttctggaactcttccccaaggggtcacttctgtcaggaagag 1020
DB 961 GACAGCAACAGATCTTCTTGTGGCACTCTTCCCAAGGGGTCACTTGTGTGTCAGGACAG 1020

QY 1021 ccagcctagcactcctctgggcagcaacaggttcttcttggcactcttccccaggatga 1080
DB 1021 CCAGCGCTTAGCACTCCTGGGAGCAACAGTCTTCTTGGCACTCTTCCCCAGAGTCA 1080
QY 1081 cttgtgtcaggaagcagccacgcttagcactcctcgtggagcaacagatcttctgga 1140
DB 1081 CTTTGTGTGAGGAGCAGGCGCAGCCCTAGCACTCTGCGGACGACAAAGTCTTTTGGGA 1140
QY 1141 ccccttcccaagagccacttctgtcaggaatcaaggtatgctcgtgaaacatccgaattc 1200
DB 1141 CCCCTTCCCAAGAGCCACTTTTGTGTGAGATCAAGTATGCTCTGAAATCCGAGTTTC 1200
QY 1201 ctggaacagaggaacagcttagtggacagtggtgtgccccctcagaattcacaacctgt 1260
DB 1201 CTGGACAGCAGGAGACAGTCTAGTGGACAGTGTGTGGCCCCCTCAGATTCACAAACCTGTCA 1260
QY 1261 tgcgcacatccatgaacatgaatggaatgggtatccctctctcaccgatacagccaa 1320
DB 1261 TGCGGCATCCATGAACATTTAGTGAATGGGATCTCTTCTACCGCATACAGCCCA 1320
QY 1321 acctggatcacctggagatctctgttggacatacagaagacgggggagacatctcgagc 1380
DB 1321 ACCTGGATCAGCTGGATCTCTGTGTGACATACAGAGCGGGGAGCATCTCGAGGC 1380
QY 1381 aaaaaccgagagagtcgcgcacgcatgcatggaagcctgtgcaacacacacaa 1440
DB 1381 AAAAAACCGAGAAGTCCGCGACAGCATGATGCCATGAGGCTGTGCAACACCAACCA 1440
QY 1441 cgggtattataggaacagcagatctctgtgtgacacaggaatgtcaaatgaat 1500
DB 1441 CGGTATTTATAGGCAACACGAGTATCTTGTGACACAGAAATGTAACATTAATTAAT 1500
QY 1501 ctccatgcaaacactgctgcacatgctgctgaagctaaataaacagagttatcttct 1560
DB 1501 CTCATGCAACACTGCTGTGCCAATGCTCTAGCGCTAAATTAACAGAGTTTAATCTTCT 1560
QY 1561 ttctctcgctcttcttggaaagttggaatgggaatggaagctatggtcgtgaatga 1620
DB 1561 TTTCTCTCGCTTTTGAAGCTTGGATGGGATGGAGTGTATGGGCTGAAGTGAAT 1620
QY 1621 gccaacacttctcttgcacaaagacacatcggaatcttctcgttggcaggaatttg 1680
DB 1621 GCCAACACTTCTTCTGCCCCAAGAACACATTCGAGTGTCTCTGCGCAGAGATTTGG 1680
QY 1681 taacagatctcccgagagattagacagccttgagatcacatgatgaatcagttat 1740
DB 1681 TAAACAGATTTCCCGAGATTTAGAGACGCTTGGAGTACCATGATTAACAGATTTAAAC 1740
QY 1741 ttctcaattatttattcttctgtttatatacccgagccaactgaaagaatgctc 1800
DB 1741 TTCTCAATTATTATTATCTTCTGTATTATCCGACCAATCTGAGAAATAGCTC 1800
QY 1801 gaattcaagctcccttagaagatgtggatc 1830
DB 1801 GAATTCAAGCTCCCTTAGAAGTGTGGATC 1830

RESULT 2
ECUS666
LOCUS 1831 bp DNA linear INV 04-SEP-1998
DEFINITION Encephalitozoon cuniculi gene encoding polar tube protein.
ACCESSION AJ005666
VERSION AJ005666.1 GI:3451140
KEYWORDS polar tube protein.
SOURCE Encephalitozoon cuniculi.
ORGANISM Encephalitozoon cuniculi
Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
REFERENCE Delbac F.
AUTHORS Direct Submission
TITLE Submitted (21-Apr-1998) Delbac F., Universite Blaise Pascal UPESA
JOURNAL CNRS 6023, Laboratoire de Biologie Composee des Protistes, 24

REFERENCE avenue des Landais, 63177 Aubliere, FRANCE
2 (bases 1 to 1831)
AUTHORS Delbac, F., Peyret, P., Metenier, G., David, D., Danchin, A. and
Vivares, C.P.
TITLE On proteins of the microsporidian invasive apparatus: complete
sequence of a polar tube protein of Encephalitozoon cuniculi
JOURNAL Mol. Microbiol. 29 (3), 825-834 (1998)
MEDLINE 98389314
FEATURES
source Location/Qualifiers
1. 1831
/organism="Encephalitozoon cuniculi"
/db_xref="taxon:6035"
sig_peptide 345..410
CDS 345..1532
/codon_start=1
/product="polar tube protein"
/protein_id="CA06662.1"
/db_xref="GI:3451141"
/db_xref="SPTREMBL:076942"
/translation="MKGISKILSASIALMKLENYSATALCSNAYGLTPGQGNAAQPP
SYVLIPSTPTIANCASGSDITYSPSPAATSPVTPKTSSENETSPAPADYGTCKI
AVLKHCDAPCTTGTTGSGPCEPPEQOOLSYISTPAPVTVESASQSPVVPVY
VAHNOAVPGYNNGTSGIPGQOOLISGLTPRPAATLCOGMAPSPGQOOLISGLTPG
VTHCOQAPSTSPGQOOLISGLTPRPAATLCOGMAPSPGQOOLISGLTPGATLCD
QGMPTSGVPGQOOLISGLTPRPAATLCOGMAPSPGQOOLISGLTPGATLCD
DIQGTGTSCEQRPERSATQYAMEACATPTPIYIIGNSEYLVPGMNAINSPONTAV
QCC"
mat_peptide 411..1529
BASE COUNT 460 a 487 c 434 g 450 t
ORIGIN
Query Match 100.0%; Score 1830; DB 3; Length 1831;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gaattcagatgctcattacattggaattaaataatgtagttcatttggatataatcctg 60
Db 1 GAATTGAGATGCCATCAATCTTGCGATTAAATAATGATGTTCAATTTGTTATATCTCG 60
QY 61 ggaggagagcgagcctgcttctcaggaggtgctgctaccagtgagacggaggtcc 120
Db 61 GGGGAGACGGCGGCTGTATCTTCAAGGGGTGCTGCTACCCAGTGACAGAGGTTCC 120
QY 121 ggaaggtctctgtagtgaagtaagggcatttgggtgttcattcatcattgcatgctcct 180
Db 121 GGAAGGTCTCTGATGGAAGTAAGGCCATTGTGGGTTCTCATCCATGTCATGCTCCT 180
QY 181 ttgggtgtttcaccaagatcaattatctcctcaggagcttcaaccctcagaatgaa 240
Db 181 TTGCGGTCTTCCACCAAGATCCAAATTAATCTTCCACGACTTCAACCTCAGAAATGAAA 240
QY 241 cagagatgaagactctgtgtaaatgtagatcagatggaagacattgaaaccaagag 300
Db 241 CAGAGATGAAGACTCTGTGTAATAATCGATATGATGGAAGACATTGAAACACCGAG 300
QY 301 ttctgaataaagataaatacctcgcgaataacgagagttcaagaatgaagatatttcta 360
Db 301 TTTGAATAAAGATATAAATCTCTCGAATAACGAGAGTTTAAGATGAAGATATTCTTA 360
QY 361 agatctctctgctctattgacctgtagtaagttggaagatgctcattcagcaaccgac 420
Db 361 AGATCTCTCTGCTCTATTGCTCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
QY 421 tgttcagcaatgcatatgagcctaactccggagacaacaggagtagtctagaagccgtgt 480
Db 421 TGTTCAGCAATGCAATAGGCTTAATCCGGGACAAACAGGATAGGCTAGAGCCGTGT 480
QY 481 atgtctgatacccaagaccccggaacacatagcaaacctgtgcaagcggttccacagaca 540
Db 481 ATGTCTGATCCCAAGACCCCGGGAACCAATAGCAAAATGTGCAAGCGGTTCCACAGACA 540

QY 541 caattctccttcccgctgacaccacatctccagttgactccggaggaatactagcgaga 600
Db 541 CATATTCTCTTCCCTCCGCTGACACCATCTCCAGTACTCCGGGAAACTAGCGAGA 600
QY 601 atggagacttcccatctgctctcctcagagaagtgtatgaagacttgcaagttccgtatga 660
Db 601 ATGAGACTCTCTCCAGGCTCTCTGAGAAAGTATGAGAAATGCAAGATTCCCGATTGA 660
QY 661 agcactcgacgcaccagaacacaatacagggaagcaccaggggtcaggccttctgaaa 720
Db 661 AGCACTGGAGGACGACACAGAAACAACTAGGAGACACACAGGCTCAGGCTTGTGAAA 720
QY 721 cccagagacagcaaacagccttctgtaagtatctccaccactctgctgtaacggtgactg 780
Db 721 CCCAGAGACAGCAACAGCCTTGTGATGATGATCCACCACTCTCCCTACCGGTACTG 780
QY 781 tggagctcgacagcttccatctctgtgtgccaagtgttccctgtgtgtcttaccacagg 840
Db 781 TGGAGCTCGACAGCTCTCATCTGTGTGCAAGTGTCTGTGCTTCTCACCAACAGG 840
QY 841 caattccagctactacacaataatgaacatccggtattccttggagacagacagatcctt 900
Db 841 CAGTTCCAGGCTACTACAAACAATGAAATCGGATTTCTTGACAGACAGACAGATCTTT 900
QY 901 ctgggactcttcccccaggaagccaattgtgtcagggaacaggccatgcttagcactctg 960
Db 901 CTGGGACTCTTCCCCAGAGGCACTTGTGTGACAGGAGGAGCCATGCTAGCACTCTG 960
QY 961 gacagcaacaaatccttctctgcaactcttccccaggggttaacttgtgtcaaggacagg 1020
Db 961 GACAGCAACAAATCTTCTTGACACTCTTCCCGGGGTCACTTGTGTGACAGGACAGG 1020
QY 1021 ccaagccttagcactctctggagcagcagagctccttctgagcacttcccccaggaagta 1080
Db 1021 CCAGGCTTAGCACTCTCTGGAGCAGACAGATCTTCTTGACACTCTTCCCGAGAGTCA 1080
QY 1081 ctgtgtcagaggaagcagccagccttagcactccttggcagcaagagtlccttcttgca 1140
Db 1081 CTTGTGTGACAGGAGAGCCAGCCGCTTACACTCTTGAGGACAGCAAGAGTCTTGTGCA 1140
QY 1141 ccccttctccagggagcacttgtgtcagatcaaggtatgactgtgaaatcccgagttc 1200
Db 1141 CCTTCTTCCAGAGGACCACTTGTGTGACAGATCAAGATGATGAGACATCTCGGAGTTC 1200
QY 1201 ctggaacagaggaagcagctagtgagcagttgtggtccctcaagattccaaaccctgtgca 1260
Db 1201 CTGGAACAGAGGAGAGCTAGTGAGACAGTGTGTGCCCCCTCAAGATTCCAAACTGTGCA 1260
QY 1261 tgcgcacatcatgaaatagtggaatgggtatccttctctctacacgcatalacagccaa 1320
Db 1261 TGCCGCCATTCATGAACAATTAGTGAATGGGTATCTTCTTCAACCCATACACCCAA 1320
QY 1321 acctgtgtaactggagatccctgtgtgtgacatacagaagaacgggggggacatccctgagc 1380
Db 1321 ACCTGTGATCACTGGGATCTGTGTGACATACAAAGACGGGGGACATCTTCGAGC 1380
QY 1381 aaaaacccgaagaagtcgcgcagcagtagtgcataatgagagcctgtgcaacacaaacaa 1440
Db 1381 AAAAAACCGAAGATCCGCCACGCAAGTATGCCATGAGGCTGTGTCAACACCAACCA 1440
QY 1441 cggttattataggcaacagcagatctctgttggaacgaagtgtacaaatgaat 1500
Db 1441 CGGTATTATAGGCAACGACGAGTATCTTGTGACACAGGAATGACAAATGAATTA 1500
QY 1501 ctcaatgaacacgtgtgtccaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
Db 1501 CTCAATGAACACGTGTGTCCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
QY 1561 ttctcttgctcttcttggaagctgtgagttggaatggaagagctatagggctgaagtgaa 1620
Db 1561 TTTCTTGCTCTTCTTGGAAAGCTGTGATGGGATGAGAGAGTCTAAGGCTGAAGTAAT 1620
QY 1621 gcaaacacttctctgcccagaacacatcagatgttcttctctgttgccaggaagtttg 1680

Df	1621	GCCAACTTCTTCGCCCCAACAACATGTGGATGTTCTTCTGTGGCAGAGATTG6	1680
Oy	1681	taacagattccccaagaatttagcagcccttgaggatcacatgatgaatcaagtataaac	1740
Df	1681	TAAACAGATATCCCAGAGATTTAGCAGCGCTTGAGACCATGATGAATCACTATTAAC	1740
Oy	1741	tctctaattatttattttcttcgttttatatacccagacaactctggaagaatgcctc	1800
Df	1741	TTCTCAATAATTAATTAATTTCTGTTTTATATCCGACCACATCTGAGAATAATGCCCTC	1800
Oy	1801	gattccaagctcccccttaagaatgtyggatc	1830
Df	1801	GAATTCAAGCTCCCTTAGAAGTGTGGATC	1830
RESULT	3	CNS07EGD	220294 bp DNA linear INV 23-NOV-2001
LOCUS	CNS07EGD/C	Chromosome VI of strain GB-M1 of Encephalitozoon cuniculi (Microspora).	
DEFINITION		AL590446 GI:13560073	
ACCESSION		AL590446	
VERSION		Chromosome sequence.	
KEYWORDS		Encephalitozoon cuniculi.	
SOURCE		Encephalitozoon cuniculi	
ORGANISM		Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.	
REFERENCE		1 (bases 1 to 220294)	
AUTHORS		Peyret,P., Katinka,M.D., Duprat,S., Duffieux,F., Barbe,V., Barbazanges,W., Weissenbach,J., Saurin,W. and Vivesres,C.P.	
TITLE		Direct Submission	
JOURNAL		Submitted (05-APR-2001) to the EMBL/Genbank/DBJ databases	
REFERENCE		2 (bases 1 to 220294)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (05-APR-2001) Genoscope - Centre National de Sequencage,	
COMMENT		2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : sequef@genoscope.cns.fr Web : www.genoscope.cns.fr)	
FEATURES		Equipe de Parasitologie Moleculaire et Cellulaire, UPRES A CNRS 6033, Universite Blaise Pascal, 63177 Aubiere cedex, France (E-mail : Pierre.BEVERTE@bp.univ-dpclermont.fr; Tel : (33) 4.73.40.74.34;	
source		Fax : (33) 4.73.40.76.70)	
BASE COUNT		Location/Qualifiers	
ORIGIN		1..220294	
		/organism="Encephalitozoon cuniculi"	
		/strain="GB-M1"	
		/db_xref="taxon:6035"	
		/chromosome="VI"	
Query Match		99.9% Score 1828.4; DB 3; Length 220294;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 1829; Conservative		0; Mismatches 1; Indels 0; Gaps 0.	
Oy	1	gaattcaagctgctcataccttgagattaanaaatgatgttcatttgttatatactcg	60
Df	36787	GAATTCAGATGCCTCATACTCTGGGATTAAAAAATGATGCTCATTTGTAATATATCCG	36728
Oy	61	ggcgagcaagcgcgctgctatctctccaagggtgtcgccctaccagtgcaagaggtcc	120
Df	36727	GGCGGACAGAGCGCGCTCGATATCTTCAGGGGTGTCGCTTACCCAGAGTGACAGAGGTTCC	36668
Oy	121	ggaagtgctcttgatggaagaagaaagccattgtgtgtttctatcataatgatactgctc	180
Df	36667	GGAGGTGCTTGGATGGAAGTAAGGCCATTTGTGGCTTCATCCATGTCATGCTGCTCT	36608
Oy	181	tctgcgctgttccaagaatccaattatctcccaaggacttccaacctcagaatgaaa	240
Df	36607	TTCACTGTTTACCAAGATCCAATTAATTTCTCCAGAGACTTTCAACCCCTCAGAAATGGAAA	36548
Oy	241	cagagatgaaactctctgtgcaaatcgtatagatatcagattggaagacattgaaaccaagag	300

D	b	36547	CAGAGATGAAATCTCTGTGGCAAATCGTATGATGATTGGAGATTTGAACCAACGGAG	36488
Q	y	301	tttgaataaagaataaaataaccccgaaacgcgaagtttlaagatgaagttattcta	360
D	b	36487	TTTGAATAATTAAGTAAATATACCTCCGAAACGACAGTTTAAAGTAAAGGATATTTCGA	36428
Q	y	361	agatcctctctgctctatctgctcgtagaagtgtgagaatcttatcagcaaccgcac	420
D	b	36427	AAATCCTCTCTCTCTATTTGGCCGTGATGAAATTGGAGATATCTATTCATTCAGCAACGGAC	36368
Q	y	421	tgtgagcaatgcataatggcctcaactccggagcaacaaggtatgtgtcgaagccgtgt	480
D	b	36367	TTTGACGACATGCATATGTGCTTAATCTCCGGCAACAGGGTATGGCTCAGCAGCCGTCT	36308
Q	y	481	atgtgcgtatcccgacgaccccggggaaccatagcaaacgtgtcaagcggttcacagaca	540
D	b	36307	ATGTGCTGATGCCACGACACCCGGGAGACCATATGCAAACTGTCCAAAGCGTTCAACGACA	36248
Q	y	541	catattctctctccgctgacccaatctccagtgactccggggaaactagcaga	600
D	b	36247	CATATTCTCTCTCTCCCGGTGACCCACATCTCCAGTACTCCGGGAAATATCAGAGA	36188
Q	y	601	atgagaaatctcatatcggtctctgtcgaagaatgtaggaacatgcaaatgtgcgtattga	660
D	b	36187	ATGAGAACATCTCATATCGGTCTCTTGCAAGAAATGTAGAAATGCAAGATGGCGATTGGA	36128
Q	y	661	agcactgtcaagcagcaacgaagaacaacatcagggagcagacaaaggtcgaagcctgtgaa	720
D	b	36127	AGCAGTCGACGACCAACGAGAACATCATCGGACGACACACAGGTCAGGGCTTTGTGAAA	36068
Q	y	721	ccccagagcagcaaacgcttgtcagtgatctccaacactctccgcgtacccgtgactg	780
D	b	36067	CCCCAGAGCAGCAACAGCCTTTGTGATGATGTCCACACTCTCTCCGTACCGGTGACTG	36008
Q	y	781	tggagtctgcagaagctccatctgtttgtgcagttgttctctgtgtgtccaccag	840
D	b	36007	TGGAGTCTGCACATCTCCATCTGTGTTGTGCCAGTTGTTCTCTGTCTGTCACCAACAG	35948
Q	y	841	caattccaggtctactacaacaatgtgaacatccggtattccctgaaagcaaacagatcttt	900
D	b	35947	CAGTTTCCAGGCTACTATCAACAATGTGAACTCCGGTATCTCTGTGACAGCAACAGATCTTT	35888
Q	y	901	ctgcgaacttcccccaaggagcgaacttgtgtcagggaaagccatgctcgaactctg	960
D	b	35887	CTGGCACTTCTCCCAAGAGGCACATTTGTGTACAGGACAGGCCATGCTACACTCTCG	35828
Q	y	961	gaacgaacaagatcctttctgtgcaactcttccccaggggtcaacttgtgtcagggacag	1020
D	b	35827	GACAGCAACAGATCTCTTTCTGTGCACCTTCCCAAGGGTCACTTGTGTCAGGGACAGG	35768
Q	y	1021	caagcctctgaactctgtggcagcaaacagtgctcttcttggaacttcccccaagagta	1080
D	b	35767	CAAGCCTTAGCACTCTGTGGGACGACACAGGTCCTTTCTGTGACTTTCCCAAGAGTGA	35708
Q	y	1081	cttctgttcaaggagcagcagcgcctagaactctcgggcagcaaacaggttctcttctgca	1140
D	b	35707	CTTTGTGTACGGGACAGGCCACAGGCTAGACATCTGGGGACGACACAGTCTTTGTGGA	35648
Q	y	1141	ccctctcccaaggagcgaacttgtgtcagatcaaggtatgcctggaacatccggagttc	1200
D	b	35647	CCCTTCTCCCAAGAGGCACATTTGTGTACAGATCAAGATATCGCTGAAACATCCGAGTTTC	35588
Q	y	1201	ctggagcagaggaagatctatgtgacagtggtgtgtgccccttaagttccaaccctgtca	1260
D	b	35587	CTGGACAGGAGGACAGTCTATGTGACATGTTGTGCCCTCTAGATTCGAAACCTGTGA	35528
Q	y	1261	tgcgcacatcccatgaacattagtggaatgggtatctctctctctcagcatacagcccaa	1320
D	b	35527	TGCGCGCATCCATGAACATTATGTGAAATGGGTATCTTCTTCTACCGATACAGCCAA	35468
Q	y	1321	acctcgatcaactggtactctgtgttgacatacagaagacggggggagatcctctgcagtc	1380

Db 35467 ACCTGGATCACTGGGATCTGTGTGACATACAGAACGCGGGGGACATCTGCGAGC 35408
Qy 1381 aaaaaccgaaagtcgcgcacagatagcatggaagcctgtgcaacacacacacaa 1440
Db 35407 AAAAAACCGAAGATCCGCCACGAGATGCGCATGAGAGCCTGTGCACACCAACACCA 35348
Qy 1441 cggttatataaggaaacagagatctctgttgaccagaaatgacaaatgaaat 1500
Db 35347 CGGTATATATAGCAACAGCAGATATCTGTTGGACACAGAAATGACAAATTAAT 35288
Qy 1501 ctccatgcaaacatgctgctcaatgctgctgctgctgctgctgctgctgctgct 1560
Db 35287 CTCCATGCAACATGCTGCTCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 35228
Qy 1561 ttcttcgctctcttggaagcttgatggaagatgagatgagatgagatgagatgagat 1620
Db 35227 TTTCTTCGCTTTTGGAACTTGATGAGATGAGATGAGATGAGATGAGATGAGAT 35168
Qy 1621 gccaaactctctctgcccagaaacacatcgaatgctctctctctgctgctgctgctgct 1680
Db 35167 GCCAAACATCTCTGCTGCCAGAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 35108
Qy 1681 taagaagatcccgagagattagcagcctgagagatgacatgacatgacatgacatgac 1740
Db 35107 TTAAGAGATTCCTCCGAGATTTAGCAGCCTTGAGATGATGATGATGATGATGATGAT 35048
Qy 1741 ttctcaaatattatctctctgcttcttcttcttcttcttcttcttcttcttcttcttct 1800
Db 35047 TTCTCAATTTTATTTATTTCTTTCTTTATATCCCGAGCCAACTTGAGATGATGATGAT 34988
Qy 1801 gaattcaagctcccttaagaagtgtggaatc 1830
Db 34987 GAATTCAGCTCCCTTAGAGTGTGGATC 34958

RESULT 4
AF310677 1019 bp DNA linear INV 17-JUN-2001
LOCUS AF310677
DEFINITION Encephalitozoon cuniculi I polar tube protein gene, partial cds.
ACCESSION AF310677
VERSION AF310677.1 GI:14484923
KEYWORDS
ORGANISM Encephalitozoon cuniculi.
SOURCE Encephalitozoon cuniculi.
REFERENCE 1. (bases 1 to 1019)
AUTHORS Xiao, L., Li, L., Vivesvarea, G.S., Moura, H., Didier, E.S. and Lal, A.A.
TITLE Genotyping Encephalitozoon cuniculi by Multilocus Analyses of Genes
with Repetitive Sequences
JOURNAL J. Clin. Microbiol. 39 (6), 2248-2253 (2001)
MEDLINE 21270266
PUBMED 11376065
REFERENCE 2 (bases 1 to 1019)
AUTHORS Xiao, L., Li, L., Vivesvarea, G.S., Moura, H. and Lal, A.A.
TITLE Direct Submision
JOURNAL Submitted (04-OCT-2000) Division of Parasitic Diseases, National
Center for Infectious Diseases, Centers for Disease Control and
Prevention, 4770 Buford Highway, Atlanta, GA 30341, USA
FEATURES
source
1..1019
/organism="Encephalitozoon cuniculi"
/isolate="CDC:V385"
/specific_host="Homo sapiens"
/db_xref="taxon:6035"
/note="genotype: I"
1..>1019
/product="polar tube protein"
1..>1019
/codon_start=3
/product="polar tube protein"
/protein_id="AAK62814.1"
/db_xref="GI:14484924"
/translation="STAMKLENYSATALCSNAYGLTPEQGNAAQPSYVLIPSTPG

BASE COUNT 238 a 312 c 251 g 218 t
ORIGIN
Query Match 55.7%: Score 1019; DB 3; Length 1019;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
EORPE"
TIANCASGSDITYSPSPAPRTSPVTPKTSNETSPSPAPADYGTCKIAVLKCHDAPG
TTSGTTPGSGPETERPEOQOPLSVISTPAPVIVESAQSPVVPVVAHDAVPGY
VNNKTSIGIPGQOOLISITLPATPGATSPGQOOLISITLPATPGATPGATPGAT
STPGQOOLISITLPATPGATPGATPGATPGATPGATPGATPGATPGATPGATPGAT
GQOQSSGQCCAPDIPNPVPPSPNNISNGTSPSTAIAPNLGSLGSCVDI0KTG6T5C
373 cctcatgagcctatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 432
Db 1 CCTTATGACCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 60
Qy 433 catatgcttaactccgagaaacacagagatgagatgagatgagatgagatgagatgagat 492
Db 61 CATATGGCTTAATCTCGGGAGACAGAGGTATGCTCAGACAGCCTGATGTGATGCTCAG 120
Qy 493 ccagcaccgaggaaacatagcaaatgagatgagatgagatgagatgagatgagatgagat 552
Db 121 CCAGCACCCCGGAAACATAGCAACTGTGCAAGCGGTTCACAGACACATATTCCTT 180
Qy 553 ctccgcctgaccccaactctcagatgagatgagatgagatgagatgagatgagatgagat 612
Db 181 CTCCCGCTGACCCACATCTCCAGTACTCCGGGAAATAGCAGATGAGATGAGATGAGAT 240
Qy 613 catcgcctctcagaaagatgagatgagatgagatgagatgagatgagatgagatgagat 672
Db 241 CATCGGCTCCTGCAAGAAATGAGAACATGCAATGCGGATTAAGACATGCTGAGAG 300
Qy 673 caccagaaacacatcagagagacacacagagatgagatgagatgagatgagatgagat 732
Db 301 CACCAGAAACATCAGACGAGACGACACAGGCTCAGGCTGTGTAACCCAGACAGAGC 360
Qy 732 aacagcctctcagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 792
Db 361 ANAGCCTTTGTACATGATCTCCACCATCTCTGCGTACCGGTGACTGTGAGATGAGAC 420
Qy 793 agtctcaactgtgtgtgagatgagatgagatgagatgagatgagatgagatgagatgagat 852
Db 421 AGTCCTCATGCTGTGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 480
Qy 853 actacaacatgagacatcagatgagatgagatgagatgagatgagatgagatgagatg 912
Db 481 ACTACAAATGAGAACATCCGATATCTCGACACACAGATCTCTGCTGCTGCTGCTGCT 540
Qy 913 cccagagagcactgtgtgagatgagatgagatgagatgagatgagatgagatgagatgagat 972
Db 541 CCCAGAGACCATTTGTGTGAGAGACAGGCTCAGACATCTCTGAGACAGACAGACAG 600
Qy 973 tcccttcgacatctctcccaagagatgagatgagatgagatgagatgagatgagatgagat 1032
Db 601 TCCTTTGAGGACATCTTCCCGAGGCTCCTTTGTGAGAGACAGGCTCAGACAGGCT 660
Qy 1033 ctcttcgagcagaaagatgagatgagatgagatgagatgagatgagatgagatgagat 1092
Db 661 CTCTGAGGACAGACATGCTCTTCTGACACATCTTCCCGAGAGATCTTGTGCTCAG 720
Qy 1093 gacagagcagcctgagatgagatgagatgagatgagatgagatgagatgagatgagat 1152
Db 721 GACAGGCAAGCTAGCATCTCTGAGACAGACAGGCTCTTGTGAGACAGGCTCTCAG 780
Qy 1153 gacgcaactgtgtcagatgagatgagatgagatgagatgagatgagatgagatgagat 1212
Db 781 GAGCAGCTTTGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 840
Qy 1213 gacagatcagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1272
Db 841 GACAGCTTATGAGACAGTGTGCTGCTCAGATTCAGAACCCCTGCTATGCGGCTATCA 900

QY 1273 tgaacattgtagaattggtatcctctctacccgatacaagcccaaacctcgatcac 1332
|||||
Db 901 TGAACATTAGTGAATGGGTATCTCTCTACCGCATACAGCCCAAACTCGGATCAC 960
|||||
QY 1333 tgggataccgtgtgtgacatacagaagaaggggggagacatctcgcgagcaaaacccgag 1331
|||||
Db 961 TGGGATCTCTGTGTGACATACAGAAAGCGGGGAGCATCTCGAGCAAAAACCCGAG 1019
|||||
RESULT 5
AF310678 1018 bp DNA linear INV 17-JUN-2001
LOCUS Encephalitozoon cuniculi II polar tube protein gene, partial cds.
DEFINITION AF310678
ACCESSION AF310678.1 GI:14484925
VERSION
KEYWORDS
SOURCE Encephalitozoon cuniculi.
ORGANISM Encephalitozoon cuniculi
Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
REFERENCE 1 (bases 1 to 1018)
AUTHORS Xiao,L., Li,L., Visvesvara,G.S., Moura,H., Didier,E.S. and Lal,A.A.
TITLE Genotyping Encephalitozoon cuniculi by Multilocus Analyses of Genes
with Repetitive Sequences
J. Clin. Microbiol. 39 (6), 2248-2253 (2001)
JOURNAL
MEDLINE 21270266
PUBMED 11376065
REFERENCE 2 (bases 1 to 1018)
AUTHORS Xiao,L., Li,L., Visvesvara,G.S., Moura,H. and Lal,A.A.
TITLE Direct Submission
Submitted (04-OCT-2000) Division of Parasitic Diseases, National
Center for Infectious Diseases, Centers for Disease Control and
Prevention, 4770 Buford Highway, Atlanta, GA 30341, USA
LOCATION/Qualifiers
1. .1018
/organism="Encephalitozoon cuniculi"
/isolate="m104"
/specific_host="mouse"
/db_xref="taxon:6035"
/note="genotype: II"
<1. .>1018
/product="polar tube protein"
<1. .>1018
/codon_start=-3
/product="polar tube protein"
/protein_id="AAK62815.1"
/db_xref="GI:14484925"
/translation="TAKMKLENYSATLCSNAYGLTPGQOQMAOPSVYLIPRPGT
IANCASGSODTPSPSPAPPTSPYTPGKTSSENTSPASAPEDVGTCKIAYLAKCDAPGT
TSGITPGSPCEPTEPQOQPLSYISTTPAVPTVESKQSPVVPVVAHVHAKDAPGY
NNGTSGIPGQOQILSGTLPFGATLCOGAMPSTPGQOQVLSGTLPFGVTLCOGQAPTS
TPGQOQVLSGTLPFGVTLCOGQAPSTPGQOQVLSGTLPFGATLCOGQAMPSTPGQOQVLSGTLPFGVTLCOGQAPTS
OQSGSGCCACAPQIPNPVMPSPMISGNYPSSTAYSPMLGSLGSCVDIQRGTSCSE
OKPE"
BASE COUNT 237 a 309 c 254 g 218 t
ORIGIN
Query Match 55.3%; Score 1011.6; DB 3; Length 1018;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1014; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 556 ccgctgacccacatctccagtgactccggygnaaaactagcgaagtgaacatctccat 615
|||||
Db 181 CCCTGCAACCCACATCTCCAGTGTACTCCGGGAAACTAGGAGATGAGACTCTCCAT 240
|||||
QY 616 cggctcctgacagaatgtagaacaatgcaaatgtccglatltaagcaactgcgacgc 675
|||||
Db 241 CGCTCTCTGAGAAAGTGTAGAACATGCAAAATTCCTGATTAAGACACTCGAGCGAC 300
|||||
QY 676 caggaaacaatcaggaagacacacaggtcagggtcgtgtgaaccccaagacgaac 735
|||||
Db 301 CAGGAACACATCAGGAGACACACAGGATCAGGGCTGTGAAACCCAGACACAGAAC 360
|||||
QY 736 agccttgcagtgatctccacacactctgcgtacaggttgactgtgagctgcagct 795
|||||
Db 361 AGCCTTGTGACAGTGTCCACACACTCTGCGCTACCGGAGATGTGAGTGTGCACACT 420
|||||
QY 796 ctccatctgtgtgcaggtgtgtcctgtcgtgtgtctaccacaaagcagttccagctact 855
|||||
Db 421 CTCCATCTGTGTGCCAGTGTGCTGTGCTGTGCTCACACAGGAGTTCAGGCTACT 480
|||||
QY 856 acacaatgnaacatccgtatctccggaacagaacagatctcttgacactctccccc 915
|||||
Db 481 ACAATATGGAACATCCGGATTCCTGACAGCAACAGATCTTGTGCACTCTTCCC 540
|||||
QY 916 caggagccacttgtgtcagggagcaagccatgcctagcaactcctgagacagcaacagatcc 975
|||||
Db 541 CAGGACCACTTGTGTGACAGGACAGGACATGCTTGACATCTGTGCGACTCTTCCC 600
|||||
QY 976 ttcttgcaactcttccccaagggatcaattgtgtcagaagaaagccagcagctagcactc 1035
|||||
Db 601 TTTCTGTGCACTCTTCCCAGGGGTCACTTGTGTGTCAGGAGACAGGCTCAGCATAC 660
|||||
QY 1036 ctggagagcaacagatccttcttgacactcttccccaaggaatcattgtgtcagaagac 1095
|||||
Db 661 CTGGGAGCAACAGTCTCTTCTGSCACTCTTCCCAGAGTCACTTGTGTGCAAGGAC 720
|||||
QY 1096 aggcagcagcctagcaactctctggaacaaaggtctcttctgcaacactctccagag 1155
|||||
Db 721 AGGCACGCGCTTACCACTCTGCGGACCAACAGGTCTCTTGTGCACTCTTCCCAGAG 780
|||||
QY 1156 ccaacttgtgtcagatcaaggtatgctctggaacacacagcaggtctctggaacagagac 1215
|||||
Db 781 CCACCTTGTGTGATGATCAAGGTATCTGTGAACATCCGAGTCTGTGACACGACAGGAC 840
|||||
QY 1216 agctcagtgacaggtgtgtgccccctcagatcccaaacctgcagtcgcccacataga 1275
|||||
Db 841 AGCTGTGTGACAGTGTGTGTGCTTCAAGTTCACAAACCTGTGATCCGCCATTCATGA 900
|||||
QY 1276 acattagtgnaaattggtatcctctctctacccgatacagcccaaacctcgcgatactg 1335
|||||
Db 901 ACATATGTGGAATGGGTATCTCTTCTTACCCCATACAGCCCAAACTCGGATCACTGG 960
|||||
QY 1336 gatccctgtgtgtgacatacagaagaaggggggagacatctcgcgagcaaaacccgagaa 1393
|||||
Db 961 GATCTGTGTGTGACATACAGAAAGCGGGGAGCATCTCGAGCAAAAACCCGAGAA 1018
|||||
RESULT 6
AF310679 936 bp DNA linear INV 17-JUN-2001
LOCUS Encephalitozoon cuniculi III polar tube protein gene, partial cds.
DEFINITION AF310679
ACCESSION AF310679.1 GI:14484927
VERSION
KEYWORDS
SOURCE Encephalitozoon cuniculi.
ORGANISM Encephalitozoon cuniculi
Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
REFERENCE 1 (bases 1 to 936)
AUTHORS Xiao,L., Li,L., Visvesvara,G.S., Moura,H., Didier,E.S. and Lal,A.A.
TITLE Genotyping Encephalitozoon cuniculi by Multilocus Analyses of Genes
with Repetitive Sequences
J. Clin. Microbiol. 39 (6), 2248-2253 (2001)
JOURNAL

MEDLINE 21270266
 PUBMED 11376065
 REFERENCE 2 (bases 1 to 936)
 AUTHORS Xiao, L., Li, L., Vivasvevara, G.S., Moura, H. and Lai, A.A.
 TITLE Direct Submission
 JOURNAL Submitted (04-OCT-2000) Division of Parasitic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA 30341, USA

FEATURES
 source
 1. 936
 /organism="Encephalitozoon cuniculi"
 /isolate="CDC-V282"
 /specific_host="Homo sapiens"
 /db_xref="taxon:6035"
 /note="genotype: III"
 <1. >936
 /product="polar tube protein"
 <1. >936
 /codon_start=3
 /product="polar tube protein"
 /protein_id="AAK62816.1"
 /db_xref="GI:14484928"
 /translation="ALMKLENYVSATALCSNAYGLNPGGQGMADPSPVLIIPSTPTI
 ANCGSDPTSPSPAPMSPVMPKTSSEMTSPSPADNVGCKIAVLKHCAPRT
 SCTPSSGCEETPBOQPLSVISTPAVPTVSASPSVPPVVAHAKHAPGTYN
 NGTSGIPGQDQILISGLPGLATLCOGQAMPTSGVPGQDQISGLPGLATLCOGATPST
 PGQOVLISGLPGLATLCOGQAMPTSGVPGQDQISGLPGLATLCOGATPST
 NGYPSSTAVSPNIGSLGSCVDIQTGTGTSCQKPE"

BASE COUNT 221 a 282 c 234 g 199 t
 ORIGIN

Query Match 46.2%; Score 844.8; DB 3; Length 936;
 Best Local Similarity 92.1%; Pred. No. 5.5e-251;
 Matches 934; Conservative 0; Mismatches 2; Indels 78; Gaps 1;

QY 379 ttgcccgtgagaagttggaagaatctatcagaaccgcaatgtagcaagaatgcatg 438
 |||||
 DB 1 TTGGCTGATGAAATTGGAGAAATCTATTCTACGACCGCCTGTGACGACATGCAATG 60

QY 439 gacctactccgggacacaggtatgctcagacgcgtgtatgtctgtatcccgagca 498
 |||||
 DB 61 GCTTAATCCGGGACACAGGATATGCTCAGCCGCTGATGTGCTGATCCCAACA 120

QY 499 ccccggaaccatagcaaatgtgcaagcgttcacagagacacatctctctcccg 558
 |||||
 DB 121 CCCCAGGACATAGCAAACTGTGCAAGCGGTTCAAGACATATTTCTCTCTCCG 180

QY 559 ctgacacacatctcagatgaatcgcgggaaacatagcgaatgagacatctccatcg 618
 |||||
 DB 181 CTGCAACCAATCTCCAGTGTCTCCGGGAAACTACGAGAAATGACATCTCCATCGG 240

QY 619 ctctctgagaagaatgtaggaacatgacaaatgttcgcfatggaacatctgagacag 678
 |||||
 DB 241 CTCTCTGCAAGATGTAGCAATGCAAGATGCGGATTTGAAACACTGCGACGACACAG 300

QY 679 gaacaacatcagggaacacacaggtcagggcctgtgaaccaccagagcaacagc 738
 |||||
 DB 301 GAACAACATCAGGAGACACACAGGCTCAGGCTTGTGAAACCCAGAGAGACAGACAGC 360

QY 739 cttgtgtagatctcacaacatctctgacgtagcgtgtgagatctgacagcttc 798
 |||||
 DB 361 CTTGTGTAGATCTTCACACATCTCCGCTACCGGTGATGAGATCTGCAAGTCTC 420

QY 799 catctgtgtgcaagttgtctcgtgtgtctcacaacagaggttcacaggtctactaca 858
 |||||
 DB 421 CATCTGTGTGCCAGTTGTCTCTCTGCTTCTACACACAGGAGTTCCAGGCTACTACA 480

QY 859 acaatggaacatccggtatctctgagacagacagatctcttctgacatcttcccgag 918
 |||||
 DB 481 ACAATGGAACATCCGATATCTCTGACAGCAAGATCTTTTCTGGCACTTTCCCCAG 540

QY 919 gaggcaattgtgtcagaggacaggtcagctcagcaatctctgagacagacagatctctt 978

DB 541 GAGCCACTTGTGTACAGGACAGCCAT----- 568
 QY 979 ctggcactcttccccagagggtcaattgtgtcagggacagagccagccttagactctg 1038
 |||||
 DB 569 -----GCTTAGCACTCTCG 582

QY 1039 ggcagcaacaggtctctctgacatcttccccagagtagcattgtgtcaggaag 1098
 |||||
 DB 583 GGCAGCAACAGGTCTTCTTGCACTCTCCGCCAGGGGTACTTTGTGTAGGAGACAG 642

QY 1099 ccaagcctagcactctctgagcagcaacaggtctcttctgacacatcttccagagca 1158
 |||||
 DB 643 CCAGCGCTTAGCACCTCGGCGACGACAGGCTCTTCTGCGACCTTCTCCAGAGCA 702

QY 1159 cttgtgtcagatcagaagttatgctctggaacatccgagttccctggcagaggaagc 1218
 |||||
 DB 703 CTTTGTGTACAGATCAAGATATGCTGTGACATCCGAGTTCTGTGACAGGAGACAGT 762

QY 1219 ctatggaacagttgtgtgccccctcagatctccaaccctgtcatgtcgccatcactaga 1278
 |||||
 DB 763 CTACTGACAGATGTTGTGCCCCCTCAGATTCACAAACCTGTGATCGCGCATCCATGAACA 822

QY 1279 ttatgtgaaatgggtatccctctctctcagcacaacagccaactctgactggat 1338
 |||||
 DB 823 TTACTGMAATGGTATCTTCTTACCCGATACAGCCCAACCTGAGATCACTGGGAT 882

QY 1339 cctgtgtgacatcagaagaaggggggagacatctctgagcaaaacccgaga 1392
 |||||
 DB 883 CCTGTGTGACATACAGAGGAGGGGAGACATCTGTGACAAACCCGAGA 936

RESULT 7
 AX007051 1116 bp DNA linear PAT 06-SEP-2000
 LOCUS AX007051
 DEFINITION Sequence 3 from Patent W00001724.
 ACCESSION AX007051
 VERSION AX007051.1 GI:9994977.
 KEYWORDS
 SOURCE Encephalitozoon intestinalis.
 ORGANISM Encephalitozoon intestinalis.
 Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 REFERENCE 1 (bases 1 to 1116)
 VIVARES, C., DANCHIN, A. and DELBAC, F.
 AUTHORS Microsporidium polar tube proteins, nucleic acids coding for said
 TITLE proteins and their uses
 JOURNAL Patent: WO 0001724-A 3 13-JAN-2000.
 VIVARES CHRISTIAN (FR); DANCHIN ANTOINE (FR); DELBAC FREDERIC (FR);
 CENTRE NAT RECH SCIENT (FR)
 FEATURES
 source
 1. 1116
 /organism="Encephalitozoon intestinalis"
 /db_xref="taxon:58839"

BASE COUNT 300 a 332 c 230 g 254 t
 ORIGIN

Query Match 9.6%; Score 176; DB 6; Length 1116;
 Best Local Similarity 53.5%; Pred. No. 6.2e-43;
 Matches 635; Conservative 0; Mismatches 460; Indels 93; Gaps 9;

QY 345 atgaagaatattcagaatcctctgacatcttccccagagtagcattgtgtcaggaagtc 404
 |||||
 DB 1 ATGAAGATATTCTAAGTCTCTGACGCTCTATGCTTGTGCTTAATGAAGTTGAAGGTGTC 60

QY 405 tattcagaacagcagctgtgacgaatgcatatggtcctaactccgggacacagaggtatg 464
 |||||
 DB 61 TATTCTAACACTGCTGCTGTGAGATTC-----AACACAAAGACTCAGAGCA-- 109

QY 465 gctcagaagcgtgtgtgtatctccagaccccggaacacatagcaaatgtgca 524
 |||||
 DB 110 -CAACCAACCGTATATGTGCTGTCTTCTAGTGCACACAGACAAATGACCAACTGTGGA 168


```

OY 525 agcgggtcacagacacatatctctctcccgctgcagccacatctccagtgactccg 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 TAGAGTCACAGAACATGATGTCCTTACTCTACTACATGCTCTCCACAGTGC 228
OY 585 gggaaacacagagagatgagatct-----ccatcggtctctgcaagaatgca 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 GGCACACTGGTGTAGAGCAGACACTCTCTCCAAACATCACTCTCTACAGAGATG 288
OY 636 ggaacatgaagatgcgtatctgaagacactgcagacacacgaacacatcaggagc 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 GGAACATGCAAGTGTCTGTGTAAAGCATTTGATGACACAGAACATCATCAACAC 348
OY 696 acacagaggtcagggcctgtgtaaacccagagcagaacacgcttctgtaagtctcc 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 TGGCAACCGGAACAGACTTGGCCCTCT-----CAGCCAGTAGACGCTACAAAT 399
OY 756 accactcccgctgaccgtgtagctgtgagctgcacagcttccatctgtgtgcaglt 815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 400 GCCACACACTGGTGTGTCTGTGTGACAGCCGCCAAGACAGCTGTACCATCTTACT 459
OY 816 gtctcgtcgtgctcacacacagcaggttccagcttactacacaaatggaatccggt 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 460 CCAAGAGCCGCTCTCTGCCAGCCGCGCAACCATC-----ATTTCCTCATTCACAGGAC 515
OY 876 attcctggaacagaacacagatctcttggaactcttcccccagagcacttgtgtag 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 516 AGGCTACTACTAGTAGCAATTCCTCGGCAAAATACAGTAAGTAATGTTCTCTCCAA 575
OY 936 ggaacagacatgctgacactctctgacagcaacagatcttcttgacacttctcccca 995
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 576 TGCCTCTTCTGCAAGTGTGTGCCGAAACAGAGAAAGCTCCACCCCGACAGAGCTAC 635
OY 996 ggggttactctgtgtaaggagcagcagcactgacacttctgagcagcaacaggtcct 1055
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 636 AGCGCGTGTTCATCTGGAACATTCCTTGGCAAAATAGTACAGGAACCTAAAG----- 689
OY 1056 tctggaactcttcccccagagtagcacttgtgtcagggagcagcagccttagcactc 1115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 690 --TAGCGGAACACACCCCTGACAGCAATCTTCCCGGGAATCTCTCTGTGGAGACCT 747
OY 1116 gggcagcaacaggtctcttctgacaccttctccagggagcacttgtgtgtagatcaa 1175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 748 CAGCCGGATCAAGCTACTCTGGAACCCCTACCC-----782
OY 1176 ggtatgcctggaacatccgaggttctctgacagcagcagagtagtagtgcaggtgt 1235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 783 -----TTCTGTATGCCAAAGCCAACTGTGGACAGCA-----ATCTCTCTGC 822
OY 1236 gcccctcagatccaaacctgtcatgacgcagcatcattgagaaatgaggtat 1295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 823 ACTCCCTCCAAATCAAAACCTGTAAATGCCAACTCTATGGGTATGAGCATATGGGTAT 882
OY 1296 ccttcttaccgcatcacagcccaaacctcagatcacttggatccctgtgtgtacacag 1355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 883 CCCAGCTCAACTGCTGTAACGCCCAACCTTGGACAAATGGGACCTTGCATCGACACAG 942
OY 1356 aagaaggggggagacatcctgctgagcaaaaaacccagagagctgcagcagctgtgc 1415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 943 AAGTC-----AACATCATCTCTGCAACCAAAAGAAAGAGCTTACACACAGTATGAA 966
OY 1416 gaggctctgtgacacacacacacacagcttatattaggaacagcagtagtctgtgga 1475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 997 GAAGCATGTGGCTGACCAACTCTCAACTGCTGTAGGAAATGCTAGATCTCTTAGC 1056
OY 1476 ccaaggaattgacaatgcaatctcattcattgcaacacagctgtcaca 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1057 CCGGGATGTATATTTACTCAACTCTCCATGCAACGCTTGTGTCGA 1104

```

RESULT 8
 AF044915 1696 bp DNA linear INV 26-AUG-1998
 LOCUS Encephalitozoon hellem polar tube protein ppp55 precursor (ppp55)
 DEFINITION

```

gene, complete cds.
ACCESSION AF044915
VERSION AF044915.1 GI:3452284
KEYWORDS Encephalitozoon hellem.
SOURCE Encephalitozoon hellem.
ORGANISM Encephalitozoon hellem.
REFERENCE 1 (bases 1 to 1696)
AUTHORS Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
TITLE Keohane,E.M., Orr,G.A., Zhang,H.S., Takvorian,P.M., Call,A.,
The molecular characterization of the major polar tube protein gene
from Encephalitozoon hellem, a microsporidian parasite of humans
JOURNAL Mol. Biochem. Parasitol. 94 (2), 227-236 (1998)
MEDLINE 98418769
REFERENCE 2 (bases 1 to 1696)
AUTHORS Weiss,L.M. and Keohane,E.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1998) Pathology/Medicine, Albert Einstein College
of Medicine, 1300 Morris Park Avenue, Room 504 Forchheimer, Bronx,
NY 10461, USA
FEATURES
source Location/Qualifiers
1..1696
/organism="Encephalitozoon hellem"
/db_xref="taxon:27973"
124..189
/gene="ppp55"
/feature="ppp55"
/feature="putative"
124..1485
/gene="ppp55"
/feature="ppp55"
/feature="N-terminal sequenced peptide was
AVPLCSNTYDPSQ"
/codon_start=1
/product="polar tube protein ppp55 precursor"
/protein_id="AAC32812.1"
/db_xref="GI:3452285"
/translacion="MKGISKILISASIMWKIGCNYSAPVPLCSNTYDPSQSPSYVLI
SPEPATITNCASPKNAVYSSPTSSSTPGNNNENSTPTEDVCKISVKKHCDTP
GASSTPCBEQIIPAOPTMATVTPALIASVOTPSVAVIVTQKVIOPAMIVPSS
IIPGYPNSTPAPAPGOGOIISGLVASGSCOLVPGNTFGOMLPGMTPGVSLPLPT
OGDGSNOTIPIVYPCOPGOGSSGSGNITIGVSPCOGOGSSGSGNITIGVSPCO
PGOGSSGSGNITIPVYPCOPGOGSSGSGNITIPVYPCOPGOGSSGSGNITIPVYSP
COGOGSSGSGNITIPVYPCOPGOGSSGSGNITIPVYPCOPGOGSSGSGNITIPVYSP
IDVQKPTSSCSQTNKSTMOYAMECAAPPTVYIGNSEYLVGMYSLTSPCNSC
CQC"
124..1485
/gene="ppp55"
190..1482
/gene="ppp55"
/feature="ppp55"
/feature="putative"
/product="polar tube protein ppp55"
BASE COUNT 460 a 433 c 373 g 430 t
ORIGIN
Query Match 7.9%; Score 144.2; DB 3; Length 1696;
Best Local Similarity 66.0%; Pred. No. 5; 3e-33;
Matches 241; Conservative 0; Mismatches 118; Indels 6; Gaps 2;
OY 1212 gacagctcagtggaacaggtgtgtccctcagatccaaacc---tgtatgcgcga 1268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1165 GGCACGCTGGGAGTGCCTCTCTCTCAAACTCCAAACCTTATTCATGCAATGCTCCA 1224
OY 1269 tccatgaacattagtggaatggtatctcttcttaccagcatacagccaaacctcgga 1328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1225 ATTTCAGGTATCATGTGAAACGGGTATCTCTACTTACTACATACACTAAAGCCTTGG 1284
OY 1329 tcaatggaatcctgtgtgtacacagaagcaggggagacatctctgcagacaaaccc 1388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1285 CAGCTGGTCTTGTGATGTGACGTAACAAATCTTCATGTGTAATCACAGAGAAAT 1344
OY 1389 gagaagctccgacagcagatgcatgtagagcgtctgtgcaaacacacacacaggtatt 1448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1345 GAAGACTTACTATGACATGACATGAGGCTTGTGCGCACACCAACCCCAACAGTTGTC 1404

```


Oy	1449	ataggcaacagcgagtatctgttgacacggaatgtaacgaatcaactctcatgc	1508
Db	1405	ATAGGTAACAGGAGATATCTGTGCGGGCGGGGAGTACAGTCACTTACCTCCATGC	1464
Oy	1509	aacactgtgtgccaatgctgtcgtacgataataaagagtttaactctcttcctcg	1568
Db	1465	AACTCTCTCTGTCAA--TGTTAGGCTAAATAAATGAGTTAACTCTTACTCTTAA	1521
Oy	1569	gtctt	1573
Db	1522	GTTTT	1526

RESULT		9				
AF338363						
LOCUS	AE338363	1340 bp	DNA	linear	INV 18-TUN-2001	
DEFINITION	Encephalitozoon hellem	genotype 2B	polar tube protein gene,	partial cds.		
ACCESSION	AF338363					
VERSION	AF338363.1	GI:14485652				
KEYWORDS	.					
SOURCE	Encephalitozoon hellem.					
ORGANISM	Encephalitozoon hellem					

REFERENCE	AUTHORS	TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS	TITLE
1 (bases 1 to 1340)	Xiao, L., Li, L., Moura, H., Sulaiman, I., Lal, A.A., Gatti, S., Scaglia, M., Didier, E.S. and Valsavara, G.S.	Genotyping <i>Encephalitozoon hellem</i> isolates by Analysis of the Polar Tube Protein Gene	J. Clin. Microbiol.	39 (6), 2191-2196 (2001)	2 (bases 1 to 1340)	Xiao, L., Li, L., Moura, H., Lal, A.A., Gatti, S., Scaglia, M. and Valsavara, G.S.	

JOURNAL
Submitted (19-JAN-2001) Division of Parasitic Diseases, Centers for
Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA
30341, USA

FEATURES	source	location/Qualifiers
		1..1340
		/organism="Encephalitozoon hellem"
		/isolate="CDC:V261"
		/specific_host="Homo sapiens"
		/db_xref="taxon:27973"
		/note="genotype: 2B"
mRNA		<1..>1340
CDS		/product="polar tube protein"
		<1..>1340

```

/product="polar tube protein"
/protein_id="NAK63046.1"
/db_xref="GI:14485653"
/translaction="MKLGNYSAVPLCSNTYDPSQQDPYSYLVSTPEATINCAYSPC
NAVYSPSTSSSTPGCTNDNETSPTEEDCTCKISYVHKHDFGASSTPCEPQTHIE
AQTMTATVTPALIASYVTPSPVAVIPYTKQVIOBATVMTIPSSILPNAMNIIDGVPYR
GTPAAPCGQGGIILSGSVLAPBASSCOLYPRGNTPQMLPGMTGYSPCLPTPGGCGSNQ
TTEIVTVCOPGQNGDSNQTIPGIVTVCOPGQNGSGSGSNQTIIPGIVTVCOPGQNGS
GSGSNQTIIPGIVTVCOPGQNGSGSGSNQTIIPGIVTVCOPGQNGSGSGSNQTIIPGIV
PCPGNSGSGSNQTIIPGIVTVCOPGQNGSGSGSNQTIIPGIVTVCOPGQNGSGSGSNQTI
GQCSYVPQTPHPIAMPPIPSGISGSGYPISTTYTSLQGLGFCIDVQKPTSTCES"

```

Query Match	5.6%	Score 102.6	DB 3	Length 1340
Best Local Similarity	66.9%	Pred. No. 4.4e-20		
Matches 204	Conservative 0	Mismatches 74	Indels 27	Gaps 3

OY 385 tatgaagttgtagaatctatcaccagaacccgacitgtgcaccaatcatgatggccaa 444
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TATGAACTGGCGAATCTTTATTCTGCAGTTCGCCCTTGACACAATACTTATGATCCCT 60

[illegible]

RESULT	10				
AF338361					
LOCUS	AF338361	1184 bp	DNA	linear	INV 18-JUN-2001
DEFINITION	Encephalitozoon hellem genotype 1A polar tube protein gene, partial cds.				
ACCESSION	AF338361				
VERSION	AF338361.1	GI:14485648			
KEYWORDS					
SOURCE	Encephalitozoon hellem.				
ORGANISM	Encephalitozoon hellem.				
REFERENCE	Eukaryota, Microsporidia; Unikaryonidae; Encephalitozoon.				
	1 (bases 1 to 1184)				

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS
Genotyping of <i>Paratuberculosis</i> and <i>Mycobacterium tuberculosis</i> H37Rv by Analysis of the Polar Tube Protein Gene	J. Clin. Microbiol.	39 (6), 2191-2196 (2001)	11376056	Xiao, L., Li, L., Moura, H., Lal, A. A., Gatti, S., Scaglia, M. and Visvesvara, G. S.
Direct Submission				
Submitted (19-JAN-2001)				
Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA 30341, USA				

FEATURES	source	Location/Qualifiers
mrna		1..1184
		/organism="Encephalitozoon hellem"
		/isolate="CDC:V274"
		/specific_host="Homo sapiens"
		/db_xref="taxon:27973"
		/note="genotype: 1A"
		<1..>1184
		/product="polar tube protein"
		<1..>1184
cds		

[illegible]

Query match 5.5%; Score 101.4; DB 3; Length 1184;

IPGVISPCQPGQGSGSNGNITPGVISPCCQPGQGNNGITGGQPGQCVSVPTPNPIAM
PPISGISGNGYPTSTTYTQSLGQLPCIDYQKPTSCESQ"

350 c	307 g	294 t
-------	-------	-------

5.48; Score 99.4; DB 3; Length 1297;

ty	66.28;	Pred. No.	4,3e-19;
ervative	0;	Mismatches	76;
		Indels	27;
		Gaps	3;

gagaatgtctatctcagcaaccgcactgtgcagcaatgcatatgacctaa 444

GGGGAATGTTTATCTGCAGTCCCGCTTGCAGTAATACTATGATCCCT 62

cagggtatgtctcagcagccgctcgtatgtgtctgataccccagcaccccg 504

-----GCACAGCAGCCATCATGTGCTGATTCGCCAGTACTCCTG 104

aactgtgcagcggttcacagacacatatctctccttcccgctgcac 564

AACTGTGCATACAGCCCAAGAATGCATATGTCCCTTCTTCTCTACTA 164

gtgactccgggaaactagcgagaattgagacatctccatcggtcctg 624 .

-----ACCCGAGGACAACAACGATACGGAACATCTC-----CTACTA 215

ggaacatgcagaattgccgtattgaagcactgcgacgcaccaggaaaca 684

AGGAACATGCAAGATTTCGGTTGTAAGCATTCGACACACACAGGTGCAT 275

1232 bp	DNA	linear	INV	11-JUN-2001
helle	protein	gene	partial	cds

1 GI:14345998

tozoon hellem.

tozoon hellem
; Microsporidia; Unikaryonidae; Encephalitozoon.

11 to 1232)
Ll, L., Moura, H., Sulaiman, I., Lal, A.A., Gatti, S.,

1., Didier, E.S. and Visvesvara, G.S.
Encephalitozoon hellem Isolates by Analysis of the polar

Microbiol. 39 (6), 2191-2196 (2001)

1 to 1232)

LI, L. / Moura, H., Lal, A.A., Gatti, S., Scaglia, M. and a, G.S.

Division of Parasitic Diseases, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA 30341 (22-JAN-2001)

Education/Qualifiers
Control and Prevention, 4/70 BUIDO highway, Altamira, CA

```

1. .1232
2004/04/01
jordanism="Encephalitozoon hellem"

```

```

/ isolate="LEPV-2-93"
/ specific host="human"

```

```

/db_xref="taxon:27973"
note="genotype: 1B"

```

```
<1. .>1232
/product="polar tube protein"
```

```
<1. .>1232
/codon_start=2
```

```
/product="polar tube protein"  
/protein_id="AAK07732.1"
```

```

/ab_xref="GI:14345999"

```

	BASE COUNT	325 A	337 C	290 G	280 T		
	ORIGIN						
	Query Match	5.4%	Score 99:	DB 3;	Length 1332;		
	Best Local Similarity	66.3%;	Prod. No.	5.6e-19;			
	Matches 201; Conservative	0;	Mismatches 75;	Indels 27;	Gaps 3;		
OY	387 atgaagcttgagaataatgctattcaggcaaacgcacgtgcagcaatgatcatgacttaact	446					
Db	2 ATGAAGCTGGGAAAGTTATTCTCAGTTCGCCTTGCAAGTAATCATTAATGATCCCTC-	60					
OY	447 ccgggacaacagggtatgctgcagcacgcctgcatgtgctgatccaccagcaccgggga	506					
Db	61 -----GCCAACGACGACATCATATGTGCGCATTCGCCAGTACTCTTAG	103					
OY	507 accatacgaaactgtgcagaagcggttcacagaacatatctccttcccgctgaccc	566					
Db	104 GCTATAACAACACTGTGCATACAGCCCCAAGAATGCAATGTGCCCTTCTCCCTACTACA	163					
OY	567 acattccaagtaccccggggaaaacctgagcaatgatgaactctccatcgctctctga	626					
Db	164 TCGCTTCA--ACCCCAGGGACAAACAACGATTAACGAACACTCTC-----CTACTACA	214					
OY	627 gaagatgtaggaacatgcagaatgtcgctatgaagacactgagcagcacccaggaacaca	686					
Db	215 GAGGATGTATGAGAACATGCATGCAAGATTCCCTGTGTAAGCATTTGGCACACACAGTGCATCA	274					
OY	687 tca 689						
Db	275 TCA 277						
RESULT 13							
AF429315/C	125020 bp	DNA	linear	PRI 18-JAN-2002			
LOCUS							
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.						
ACCESSION	AF429315						
VERSION	AF429315.1	GI:17646244					
KEYWORDS							
SOURCE	human.						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1 (bases 1 to 125020)						
AUTHORS	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L. A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2						
TITLE	Nat. Genet. 29 (4), 377-378 (2001)						
JOURNAL	Natl. Genet. 29 (4), 377-378 (2001)						
MEDLINE	21583737						
PUBMED	11694876						
REFERENCE	2 (bases 1 to 125020)						
AUTHORS	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L. Direct Submision						
TITLE	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA						
JOURNAL	Location/Qualifiers						
FEATURES	1..125020						
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16q24.3: between D16S520 and WT-12410" /note="Isolated from a patient with Huntington's						

Repeat_Region	Disease-Like 2 (HDL2)*
repeat_region	complement(35581..35746)
mRNA	/rpt_type= tandem /rpt_unit=ctg complement(<35507..>36887)
gene	/gene="JPH3" /product="junctionhilin 3" complement(<35507..>36887) /gene="JPH3" /note="JPH3"
CDS	complement(<35507..36887) /gene="JPH3" /note="component of the junctional complex between plasma membrane and endoplasmic reticulum" /codon_start=1 /product="junctionhilin 3" /protein_id="AAL00941.1" /db_xref="GI:17646245" /translation="MSSGGRENFDDGSGYCGGWEDGAHGHGCTGPKGGEYGTGNS HGFEGLVYTPSGNTGYGTGTAQCRHIGIGLESKGVVYKKEWTHGFKRGYVECGAG NGARVEGWSNGLDGGYCTEYSFG"
BASE COUNT	29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN	
Query Match .	2.7%: Score 49.4; DB 9; Length 125020;
Best Local Similarity 12.2%: Pred. No. 0.0043;	
Matches 121; Conservative 424; Mismatches 442; Indels 8; Gaps 3;	
525 agcggltcacagacacatctcccttcctccggtgcgaaccacacatctccagtacccg 584	
Db 51634 MKSRMTTGRMRGTRRAMMMWMMKKYRAAMKCAKTRSMWMAKSKSYARKCCSMM 51575	
585 gggaaactcagagagaatgagacatctcatctgctctctgcgagaagtgtagaacatgc 644	
Db 51574 AYKKYKSRAISRCKGNGKMAKSCMRRYMMYKMGGGAGKCCSAGAMGTGRSMAOCMKMS 51515	
645 aagattgcgcatltaagacatcgagcagcagaacaacatcagaggagcaccagg 704	
Db 51514 KKSASKARAYASTGAAMSMCGTCYMKKCTRRMAATACAMRRWAMMMSCMRCKWTSK 51455	
705 tcaaggccttgtgaacccacagacagacacagccttgtca-gtgcattccaccatcc 763	
Db 51454 MSKMGRMRMYCWSMARMMYVAGSYMSKCRSRMGKGRSRSRSGMSWTRGVSYSARCYM 51395	
764 tgcgcgtaccggtgagctgtgagctgcgcagctccatctcgtctgtgcgaagtgtccctgt 823	
Db 51394 SKSAGAGKCKKAGSYAKKSAGCMRWGSYGWMGYSCRSMAKSKSRMTCTWSCCTCSKGC 51335	
824 cgttcctcacacacagcagcagctcagcgtcactcaacaatgaaatcgcgtatccctcg 883	
Db 51334 YMMYMSMKTAGCMCYMRRCCTCMARMMRAMMAARRRMAAGARAWMSMMMSMMMR 51275	
884 acagaacaagatccttctgtgcacttccccacaggaacacacttgtlcaaggagaagc 943	
Db 51274 BHAAMAMAMMAAAMCMWMTYRKYCTCMWSMAASMAAMMMWMCRAAWYKKTMTMRMS 51215	
944 catgcctagcactcctgtgacagacaagatcc---ttctgtgcacttccccagagg 999	
Db 51214 KYSSSYAAYMYMTWMTWCTYSSMYAARBMGTGSGKRRKMAWSTSCGCCRCGCMKKGR 51155	
1000 tcacttgtgtcaggaacagcgaagcctcagacactcctgtgaggaacaagtccttctg 1059	
Db 51154 WGSTTTGASGKSRMWYSSMACMTWYKGSRRKKTSMATSGCMRRGAMGSKGMKRYVSC 51095	
1060 gacacttccccacagagatcacttgtgtgtcgaaggacagccagcagctagcactcctggc 1119	
Db 51094 RFXGMRWGWMSYKCGSGYMSMGMTGSGRTSYCYCTGTGYSCKAKKCKSKSTSKSCT 51035	
1120 agcaacaggtccttctgtgcacacttccccagagacacacttgtgtcaagatcaagta 1179	
Db 51034 KSKYCSGMSGRSSYCTCWMGRCMKCMKSSCCSRSGCKSYKMAAGWMCRCGRGRGMSCTWTS 50975	

OY	1180	tgccgggaacatccgagatctctctgtaacgaaggaacagctctctgtaagagctgttgcgc	1239
OY	50974	SMKSTSMASYKSSASAGRGKMKCTKTCSCMYTYSMHKSSSHDBSGYTSYKRSYTWTCMKMD	50915
OY	1240	ctcagattcccaaacctctcatcgtccgcacatccatgatacatgtaggaatggtatcctt	1299
Db	50914	SMCWGVSDSHMMRDMMYDKMKRMKVSSGTGRYHBSHSMCMKSMRSRDGSGCMYHMKSYWK	50855
OY	1300	ctctacccgcatcacgccccaaacccctcgagatctgtgagctctgttcttgatcaacgaaga	1359
Db	50854	SSASHHMTYISGKSSVYBDDCMSMTTBSKSYRMRRRSGSTBYRKCKCYMRYHRFST	50795
OY	1360	cgggggggagacatcccgagcaaaaac-----cgaagaatccgcacgcagatgacatg	1416
Db	50794	RSVWYMGKRSCQVNDVSYSRRBSVHTSMSWMRCRCRGARBMKSCAGKMSCVMTGMCYHSR	50735
OY	1417	aggcctgtgtcaaaccaaccaacacggtattatattaggcaacagcagatctgttgac	1476
Db	50734	CGSMKCMSCDCCCRHADDHVCRCBMYKBDACBYKRDYSVCMSAHVWVVBHBSKS	50675
OY	1477	caggaatgtacaatgcatlaactaactcctcatgcaac	1511
Db	50674	SRBGWRHMHACHKSMCKKSVYKWHVHKMKCYWC	50640
RESULT	14		
LOCUS	AF429315	125020 bp	DNA linear PRI 18-JAN-2002
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.		
ACCESSION	AF429315		
VERSION	AF429315.1	GI:17646244	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 125020)		
AUTHORS	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,		
	Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,		
	Potter,N.T., Ross,C.A. and Margolis,R.L.		
	A repeat expansion in the gene encoding junctophilin-3 is		
	associated with Huntington disease-like 2		
	Nat. Genet. 29 (4), 377-378 (2001)		
TITLE			
JOURNAL			
MEDLINE	11583737		
PUBMED	11694876		
REFERENCE	2 (bases 1 to 125020)		
AUTHORS	Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.		
TITLE	Direct Submision		
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical		
	Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA		
FEATURES			
Source	Location/Qualifiers		
	1..125020		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="16"		
	/map="16q24.3: between D16S520 and WI-12410"		
	/note="Isolated from a patient with Huntington's		
	disease-Like 2 (HDL2)"		
	complement(3581..35746)		
	/rpt_unit=tandem		
	/rpt_unit-ctg		
	complement(<36507..>36887)		
	/gene="JPH3"		
	/product="junctophilin 3"		
	complement(<36507..>36887)		
	/gene="JPH3"		
	/note="Jp3"		
	complement(<36507..>36887)		
	/gene="JPH3"		
	/note="component of the junctional complex between plasma		
	membrane and endoplasmic reticulum"		
	/codon_start=1		
	/product="junctophilin 3"		

Query Match	Best Local Similarity	Score	DB	Length
Matches 76; Conservative 286; Mismatches 298; Indels 4; Gaps 2;	11.4%; Pred. No. 0.018;	2.6%;	9;	125020;
Db 464	ggcctagcagccgctgtagtgcctgtagtgcctcccccagcagcccccgggaagaccatagcaactgctgc	523		
Db 16900	SVSYVKHSMHMASBSCHMBKCMTWSCCMAMTKSSMMWSSWGMCCKMRKRKMKWYS	16959		
QY 524	aagcggttcacaggaacacataltctctctctccgcctgcacacatctccagtgatctac	583		
Db 16960	RGMSRSKSMRYTGGSKMMHSSMCTSSCYASMCCKMCCMRSCCCMRSCCMRYC	17019		
QY 584	gggggaacatagcaggaatgaacatctccatccgctcctctgcagaagatgtgagaacatg	643		
Db 17020	CACKCYMSSYMTWASASYMSRRSYMKRSMKCMSCRMCSRSRSKCKSGCGGSGMKKGGS	17079		
QY 644	caagatggccgattgaagcaactgcagccagcaggaacaaacataggaagacacacag	703		
Db 17080	YERKTRKRSRGM -RGAKMYMYRRSMRKMKYSSKGYCMYCMWGRRCGCTSCMTSRS	17138		
QY 704	gtcagggcctgtgaaacccacagacgaacagccttgcattgtagtattccacacatcc	763		
Db 17139	AMCGSYCAKCKSMCYSCYGMGMGYAACRYRSMSSKCYCMRSTYSTGCGCTTTT	17198		
QY 764	tgccgtaccggtgactgtgagctgcacagctccatctgtgtgccaagtgttctgt	823		
Db 17199	TCGCCCNMTGGGGAAGCTTTNCMKKITYRYRKNKGCMKCYNNYNNMSRSCRAGS	17258		
QY 824	ggtgtcaccacacagcagcttcacagctactacaacatggaacatccgattctctg	883		
Db 17259	CTRYKSSMTMASASVCWCMSTYTCMSRMSASGMSWSYMKMKMSWMSMCYKCMCK	17318		
QY 884	aaggaacagatcccttctgcacctctcccccaggaagcacttgtgtaaggagacg--	941		
Db 17319	MCTCMRSMRSGMSYMTASKSSRGCYCTRYCWMSSSCYASYVMMSKMKMKMK	17378		
QY 942	gcacatgcctagcaactcctgagacgaacaaagatccttctggaactctcccccagggt	1000		
Db 17379	WBRGMSMMGASRSRSCYKCYKSMSCMMSSKCYRACAGOMMKGGYMYMRGMSKRRMW	17438		
QY 1000	caacttgtgtcaggaagcagcagcagcttagcaactcctgtggcagcaacagtcctctg	1060		
Db 17439	GKSAMYRMRMRMKGGAMMMCKCYSRMSRCMMKSYCASCRSCAMMSGYMKASVYC	17498		
QY 1061	caactctcccccaggaagtcacttgcctgtaagggaagcagcagcctagcactccggca	1120		
Db 17499	MSSYCTCWCISYMYNCTCWCCKMSSYTKSMSSMSSTKCKRKRKSYSMCTSRGAM	17558		
QY 1121	gcaa 1124			
Db 17559	SCMR 17562			

RESULT 15

CNS05FC3

LOCUS

DEFINITION Human chromosome 14 DNA sequence BAC R-79666 of library RPCT-11 from chromosome 14 of Homo sapiens (Human), complete sequence.

ACCESSION

AL355032

KEYWORDS

HTG

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on: September 1, 2002, 00:28:16
        : Search time 303.84 Seconds
        : (without alignments)
        : 10340.816 Million cell updates/sec
```

Title: US-09-755-456-1

Sequence: 1 gaattcagatgcctcatacc.....tcccttagaagtgtggatc 1830

Scoring table:

	IDENTITY_NUC	Cancer 1 0
Cancer 10 0		

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 1008
Listing first 45 summaries

Database : N.Genseq/02802.*

1:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1830	100.0	1830	21	AA291192	Emcephallitozoon cu
2	176	9.6	1116	21	AA291194	Emcephallitozoon in
3	45.4	2.5	740	22	AAAF45087	Human secreted p
4	38.2	2.1	611	22	ABAF47116	Human breast cell
5	38.2	2.1	611	22	ABA64998	Human foetal liver
6	38.2	2.1	611	22	ABA32106	Human foetal liver
7	38.2	2.1	611	22	AAK13424	Human brain expres
8	38.2	2.1	611	22	AAK39161	Human bone marrow
9	38.2	2.1	611	22	AAI19970	Probe #9903 for ge

10	38.2	2.1	611	22	AAI05168	Probe #13654 used t
11	38.2	2.1	611	22	AAI05680	Probe #15671 used t
12	38.2	2.1	2465	11	AAO03665	Sequence homology
13	37.2	2.0	2880	10	AAO02068	Human muscarinic a
14	37.2	2.0	2880	10	AAO02068	Human muscarinic a
15	36	2.0	1090	22	AAH99405	Human protein enco
16	35.6	1.9	8269	22	AAI199642	Human expressed po
17	35.6	1.9	8269	22	ABAO6733	Human genomic DNA
18	35.6	1.9	8269	22	AAK9094.1	Human digestive sy
19	35.6	1.9	8269	22	AAI54010	Human polynucleoti
20	35.6	1.9	8269	22	AAI51472	Human DNA for a no
21	35.6	1.9	8269	22	AAI531976	Human liver associ
22	35.6	1.9	8269	22	AAI533436	DNA encoding human
23	35.6	1.9	8269	22	AAI535069	DNA #19 encoding h
24	35.6	1.9	12235	23	ABLO4214	Drosophila melanog
25	35.6	1.9	2944528	24	ABAO30341	Listeria monocytoc
26	35.4	1.9	9577	22	AAK00931	Human Immune/Haema
27	35.2	1.9	9947	23	ABLO4944	Drosophila melanog
28	35	1.9	4590	22	AAH24065	Yeast AOBp604-asso
29	34.8	1.9	3448	18	AAV06055	Singapore horsesho
30	34.8	1.9	3448	20	AAI25389	Horseshoe crab fac
31	34.8	1.9	3448	20	AAI25389	Horseshoe crab fac
32	34.8	1.9	3448	21	AAI222759	Horseshoe crab fac
33	34.8	1.9	3448	22	AAI58999	Recombinant 5' tru
34	34.8	1.9	3620	21	AAAO6581	Horseshoe crab rec
35	34.8	1.9	4182	18	AAV06054	Human cytoskeleton
36	34.8	1.9	4182	20	AAI25388	Singapore horsesho
37	34.8	1.9	4182	20	AAI25388	Horseshoe crab fac
38	34.8	1.9	4182	21	AAI222758	Horseshoe crab fac
39	34.8	1.9	4182	22	AAI58998	Horseshoe crab rec
40	34.4	1.9	2113	22	AAI58998	Horseshoe crab rec
41	34.4	1.9	10083	23	AAI888071	Human FLEXHT-2 nuc
42	34.4	1.9	10369	23	ABLI12134	Drosophila melanog
43	34.4	1.9	10369	22	AAI56303	Tumour suppressor
44	34.2	1.9	444	24	ABLI32392	Human immune syste
45	34.2	1.9	1166	20	AAK00027	Human immune/Haema
					AAI19440	M. tuberculosis an

ALIGNMENTS

PT	New microsporidial tube protein and related nucleic acid, useful e.g
DR	P-PSDB; AAT83760.
XX	
PI	Delbac F, Valvares C, Danchin A;
XX	
PA	(CNRS) CNRS CENT NAT RECH SCI.
XX	
FR	07-JUL-1998; 98FR-0008692.
XX	
PD	13-JAN-2000.
XX	
PE	06-JUL-1999; 99WO-FR01630.
XX	
OS	Encephalitozoon cuniculi.
XX	
KW	Vaccine; gene therapy; polar tube protein; microsporidium; antibody;
KW	antiparasitic; antiprotozoan; immunostimulatory; immunoassay;
XX	opportunistic pathogen; immunocompromised patient; AIDS; ds.
DE	Encephalitozoon cuniculi polar tube protein PMP55 gene.
DT	21-JUN-2000 (first entry)
AC	AA291192;
XX	
ID	AA291192 standard; DNA; 1630 BP.
RESULT	1

PT In vaccines against Encephalitozoon infection and for raising
diagnostic antibodies -
XX
PS Claim 17, Page 39-41; 55pp; French.

XX
CC This sequence represents the complete gene encoding the polar tube
CC protein PRP55 from the microsporidium Encephalitozoon cuniculi.
CC The protein has antiparasitic, antiprotozoal and immunostimulatory
CC activities. The proteins are used to raise specific antibodies, useful
CC as immunoassay reagents for detecting serum antibodies against
CC Encephalitozoon (opportunistic pathogens in immunocompromised,
CC especially acquired immune deficiency syndrome, patients), in
CC vaccines to prevent such infections.

XX
SQ Sequence 1830 BP; 460 A; 486 C; 434 G; 450 T; 0 other;

Query Match 100.0%; Score 1830; DB 21; Length 1830;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcagatgctcctacattggaatlaaaaaattgatgttcattgttatatctctg 60
DB 1 gaattcagatgctcctacattggaatlaaaaaattgatgttcattgttatatctctg 60
QY 61 ggcggacagcgcgctctcttctcaagggtgtcgcttaaccagtgacagaggtttcc 120
DB 61 ggcggacagcgcgctctcttctcaagggtgtcgcttaaccagtgacagaggtttcc 120
QY 121 ggggggtctctgtgataagaaagacattgtgtgtctctcaatctgctatcttcct 180
DB 121 ggggggtctctgtgataagaaagacattgtgtgtctctcaatctgctatcttcct 180
QY 181 ttgcgctgtttccaccaagatccaatattctctccagagctttcaacctcagatgtaa 240
DB 181 ttgcgctgtttccaccaagatccaatattctctccagagctttcaacctcagatgtaa 240
QY 241 cagaagatgaactctctgtgcaaatcgtagatatcgtattgagacattgaaacacagag 300
DB 241 cagaagatgaactctctgtgcaaatcgtagatatcgtattgagacattgaaacacagag 300
QY 301 ttggaataaaagtataaataactctccgaaaagcagaagtttaagatgaagaagatttcta 360
DB 301 ttggaataaaagtataaataactctccgaaaagcagaagtttaagatgaagaagatttcta 360
QY 361 agatcctctctgctctatctgacactccggaacagaggtatgaagatgtaattccta 420
DB 361 agatcctctctgctctatctgacactccggaacagaggtatgaagatgtaattccta 420
QY 421 tctgcaagcaatgcatatgtgcttaactccgggacacaggtatgtctcagcagcgctcgt 480
DB 421 tctgcaagcaatgcatatgtgcttaactccgggacacaggtatgtctcagcagcgctcgt 480
QY 481 atgtgtgatacccgacaccccggaacacatagcaaatctgtgaaacggttccacagaa 540
DB 481 atgtgtgatacccgacaccccggaacacatagcaaatctgtgaaacggttccacagaa 540
QY 541 catattctctctctccgctgacccacacatctccagtgactccgggaaacactagagaa 600
DB 541 catattctctctctccgctgacccacacatctccagtgactccgggaaacactagagaa 600
QY 601 atgagacatctccatcgtctctctgcaagaagtgtagaacatgcaagattgcccgtatga 660
DB 601 atgagacatctccatcgtctctctgcaagaagtgtagaacatgcaagattgcccgtatga 660
QY 661 agcaactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
DB 661 agcaactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
QY 721 ccccaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780
DB 721 ccccaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780

QY 781 tggagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
DB 781 tggagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
QY 841 cagttccagctactacaataatggaacatccgattctcttgagacagcagcagcagcagc 900
DB 841 cagttccagctactacaataatggaacatccgattctcttgagacagcagcagcagcagc 900
QY 901 ctggacactctcccccagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
DB 901 ctggacactctcccccagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
QY 961 gacagcaacagatccttctctgacactctcccccagagcagcagcagcagcagcagcagc 1020
DB 961 gacagcaacagatccttctctgacactctcccccagagcagcagcagcagcagcagcagc 1020
QY 1021 ccacgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
DB 1021 ccacgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
QY 1081 cttgtgtcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
DB 1081 cttgtgtcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
QY 1141 cccctctcccaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200
DB 1141 cccctctcccaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200
QY 1201 ctggacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
DB 1201 ctggacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
QY 1261 tgcgcacatccatgaacatagtggaatggtatctctcttccatccagcagcagcagcagc 1320
DB 1261 tgcgcacatccatgaacatagtggaatggtatctctcttccatccagcagcagcagcagc 1320
QY 1321 acctcgatcactggaatcctctgtgtgacatacagaacagcggggagacatcctgcagc 1380
DB 1321 acctcgatcactggaatcctctgtgtgacatacagaacagcggggagacatcctgcagc 1380
QY 1381 aaaaacccgagaaggtccgcgacagcagcagcagcagcagcagcagcagcagcagcagcagc 1440
DB 1381 aaaaacccgagaaggtccgcgacagcagcagcagcagcagcagcagcagcagcagcagcagc 1440
QY 1441 cggatattaagcacaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500
DB 1441 cggatattaagcacaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500
QY 1501 ctccatgcaaacactgctgtccaaatgctgtagcctaaataaaacagatttaactctct 1560
DB 1501 ctccatgcaaacactgctgtccaaatgctgtagcctaaataaaacagatttaactctct 1560
QY 1561 ttctctcgtcttttggaaagtttggaggttggaggttggaggttggaggttggaggttggag 1620
DB 1561 ttctctcgtcttttggaaagtttggaggttggaggttggaggttggaggttggaggttggag 1620
QY 1621 gccaacactctctctcccaagaacacatctcgatgttctctctggtgagcagagtttgg 1680
DB 1621 gccaacactctctctcccaagaacacatctcgatgttctctctggtgagcagagtttgg 1680
QY 1681 taacagagatcccccagagatttagcagcctgtgagatcacatgatatgtaacagatlaaac 1740
DB 1681 taacagagatcccccagagatttagcagcctgtgagatcacatgatatgtaacagatlaaac 1740
QY 1741 ttctcaattattattttcttctctgtttatattcccgagcacaatctggaagaatgcttc 1800
DB 1741 ttctcaattattattttcttctctgtttatattcccgagcacaatctggaagaatgcttc 1800
QY 1801 gaattcaagctcccttaagaagtgtgagatc 1830
DB 1801 gaattcaagctcccttaagaagtgtgagatc 1830

```
RESULT 2
AA291194
ID AA291194 standard; DNA: 1116 BP.
XX
AC AA291194:
XX
DT 21-JUN-2000 (first entry)
XX
DE Encephalitozoon intestinalis polar tube protein PTP55 gene.
XX
KM Vaccine: gene therapy; polar tube protein; microsporidium; antibody;
KW antiparasitic; antiprotozoan; immunostimulatory; immunosassay;
XX opportunistic pathogen; immunocompromised patient; AIDS; ds.
OS Encephalitozoon intestinalis.
XX
PN MO200001724-A1.
XX
PD 13-JAN-2000.
XX
PF 06-JUL-1999; 99MO-FR01630.
XX
PR 07-JUL-1998; 98PR-0008692.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Delbac F, Vivas C, Danchin A;
XX
DR MPI: 2000-237186/20.
XX
DR P-PSDB; AA183762.
XX
PT New microsporidial tube protein and related nucleic acid, useful e.g.
PT in vaccines against Encephalitozoon infection and for raising
PT diagnostic antibodies
XX
PS Claim 18: Page 43-44; 55pp; French.
XX
CC This sequence represents the complete gene encoding the polar tube
CC protein PTP55 from the microsporidium Encephalitozoon intestinalis.
CC The protein has antiparasitic, antiprotozoal and immunostimulatory
CC activities. The proteins are used to raise specific antibodies, useful
CC as immunosassay reagents for detecting serum antibodies against
CC Encephalitozoon (opportunistic pathogens in immunocompromised,
CC especially acquired immune deficiency syndrome, patients), in
CC vaccines to prevent such infections.
XX
SQ Sequence 1116 BP; 300 A; 332 C; 230 G; 254 T; 0 other;

Query Match 9.6%; Score 176; DB 21; Length 1116;
Best Local Similarity 53.5%; Pred. No. 1e-45;
Matches 635; Conservative 0; Mismatches 460; Indels 93; Gaps 9;

QY 345 atgaagaagtcttcttaagatccctctgctctatctatgcccctgtagaattggagaatgctc 404
DB 1 atgaagaagtcttcttaagatccctctgctctatctatgcccctgtagaattggagaatgctc 60
QY 405 tattcagcaacgcgactgtgcagcaatgcatatgacctaaactccgggacaacaggatg 464
DB 61 tattcacaactgtcgtgtgagatc-----aacacaaggactgagggaca-- 109
QY 465 gctcagcaacgcgactgtgcagcaatgcatatgacctaaactccgggacaacaggatg 524
DB 110 -caaccaaacgcatatgtgctgctcctagtgacacagagaataatgacaactgtgga 168
QY 525 agcggttcacagacacatatctctctcccgctgcaccacacatctccagtgactcgg 584
DB 169 tacagttccacagacacatatgtccctctactactactacatgccttcacagtgcca 228
QY 585 gggaacaactagcgagaatgagacatc-----ccatcggtctctggaagaatgta 635
DB 229 ggcaacaactgtgagagcagacacactctccacaacatcctcctacagagatgtg 288
```

```
QY 636 ggaacatgcaagatctgcgatattgaagcactgagcagccacaggaacaacatcaggagc 695
DB 289 ggaacatgcaagatgtgtgttgaaagcatgtgatgcacaggaacaacatcacaacct 348
QY 696 aacacagggtcagggccttgtgaacccacagacagcaacagccttgtcatgtatccc 755
DB 349 tgcgaaccggaaacagactttgccccctc-----cagccagttagagctatacaat 399
QY 756 accactctgcgtaaccggtgactgtgagctgtgacagatctccatctgtgtgccaagt 815
DB 400 gccacacacactgtgtgtgtctctctgtagacgcgcgaagagctgttaccatctact 459
QY 816 gtctctgtctgtccacccacagagcagttccaggtccactacacaaatggacaccggt 875
DB 460 ccaaaaggcgtctctgcacagccgcaaccatc-----attctcatcaacacggcacc 515
QY 876 attctgcaagacagacagatccttctgtgcactctccccaagagcacttgttgcag 935
DB 516 aggtactacaatagtgcaattcccggaataactacagtaatgtctctctccaaag 575
QY 936 ggaacaggcagatgcctagcactccttgacagcaagatccttcttgcaactctcccca 995
DB 576 tgcctctcttgccaagtggtgcccggaaacacaggaagctccaccccccagaccctacc 635
QY 996 gggttcacttgtgtcagggagcagcgccctagcactccttgaggcagcaagagtcctt 1055
DB 636 aggcgctgttcacatctggaacacatctccttgccaatatgtacaggaactcaag----- 689
QY 1056 tctgcaactctccccaagagatcatttgttcaaggacagcagccctagcaactcct 1115
DB 690 --tagcggaaacaccccttgacagcaatctctgcgggaatcgttctgttgaagcttc 747
QY 1116 ggagagaacagcgtcctcttgacacccctctccagagagcaacttgttgcagatcaa 1175
DB 748 cagccgatacaagctactcttgaaacccctacccc----- 782
QY 1176 ggtatgcttggaacatccgaggttccttgacagcagagagctagtgtgacagtgtgt 1235
DB 783 -----ttctgttagcaagaacatctggaacaga-----atgctgctgc 822
QY 1236 gccctcagatccaacacctgtcatatgcgcacatcatgaacatagtgaatgtgatat 1295
DB 823 actcttcacatcaaaaacccgttaattgcacactcttaagtgttcaacagatgaggtat 882
QY 1296 cctctctcaccgatacagccaaacctcgatcacttggaatctgtgtgacatcacg 1355
DB 883 cccagctcaaatgtgtgacgcccacaccttggaacaattggaaactgtcatgcacacag 942
QY 1356 aagacggggggagacatccttgagcaaaaacccgagagaagtcgcacgcaatgacatg 1415
DB 943 aagtc-----aacatcatccttggaacccaagaagaagcctgtgacacagtatggaatg 996
QY 1416 gaagcctgtgcaacacacaacaggttatattatgaagcaacagcagatcatctgttga 1475
DB 997 gaagcagtcgtgcacacacacacacatcagctgttcttaagaaatgtgtgatactcttacc 1056
QY 1476 ccaggaatgtacaatgaatcaatcaactccatgcacactgtctgcca 1523
DB 1057 ccgggagatgataatcaactcaactccatgcacagcgttgtgcca 1104

RESULT 3
AAF45087
ID AAF45087 standard; cDNA: 740 BP.
XX
AC AAF45087;
XX
DT 03-APR-2001 (first entry)
XX
DE Human secreted protein coding sequence SEQ ID NO: 26.
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
```

KW vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder
 KW neurological disease; infection; human; secreted protein; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200077023-A1.
 XX
 PD 21-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US14964.
 XX
 PR 11-JUN-1999; 99US-0138627.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis GA;
 XX
 DR WPI: 2001-032308/04.
 DR P-PDB: AAB65935.
 XX
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX
 SS Claim 1; Page 457; 558pp; English.

CC The invention relates to the isolation of genes AAF5072-745120 encoding
CC 48 human secreted proteins AAB65920-865966. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (SBRD1) for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.

Query Match	2.5%	DB 22,	Length 740;
Best Local Similarity	47.4%;	Pred No. 0.0011;	
Matches 136; Conservative	0;	Mismatches 151;	Indels 0; Gaps 0;

OY 888 ttcttcgacactcttcccccgagcgacatttgtgtcaaggagagcgactagc 957
 Db 349 ttgggggacactctcccgagcgacagcgccaactgtgtctcgaatgtcttcc 408
 OY 958 ctgagacgacaaagatccttctgtgcacttccccaggggtcactttgtcaggac 1017
 Db 409 tggccacatagtgctgtgcctcggtcaactcttctgttagtcacaggagatgtgc 468
 OY 1018 agggcaccgcttagacactccctggcgagcaaaagtctcttctgtgacatttccccagag 1077
 Db 469 agggacagctccctgtgttagcgttcaacatgatacagtcacatctctcaaccatg 528
 OY 1078 tcaatttgttcagggagacgacgacgactgacatctctgtgcagacaaagtgctttc 1137
 Db 529 gcaacatgtgtctacacatcagagacatgcacgtctgtgcacaacatgagctagtcag 588
 OY 1138 gacaccttctcccgagcgacatttgtgtcagaatcaagatgacct 1184
 Db 589 tcgaattctccagacggtcagaatgtgcttgtctaaatcaattctt 635

RESULT	4
ABA47116	
ID	ABA47116 standard; DNA; 611 BP.

DT 01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #5811.

KW Human, microarray; single exon probe; gene expression; breast, disease; cancer; ss.

OS Homo sapiens.

PN W0200157271-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 30-JUN-2000; 2000US-0608408.

PR 21-SEP-2000; 2000US-0234687.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

PT New spatially-addressable set of single exon nucleic acid probes,

PT breast, comprises number of single exon nucleic acid probes

PS Claim 4; SEQ ID NO 5811; 3277pp + sequence listing; English.
 PS
 YY

CC The invention relates to a spatially-addressable set of single exon

CC from human breast and BT 474 cells. The method involves contacting

derived from mRNA of human breast, and then measuring the label bound to each probe of the microarray. The probes are useful for

CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins that are useful for gene discovery and for

CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the likelihood of observing

CC agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression with far less bias

than expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequences.

CC present sequence is a single exon nucleic acid probe of the invention.

CC printed specification, but was obtained in electronic format directly from WIPO at <http://www.wipo.int/patdb/titles> not accessible

XX
Sequence 611 BP, 104 A, 238 C, 111 C, 150 F, 0 other.

Query Match	2.18;	Score 38.2;	DB 22;	Length 611;
Post Local 64m1jast4tur	47.39;	Prod No 0.31;		

Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

889 aacagatccttctgcaactctccccaaggagccaattgtgtcaggacagccatgc 948

Db 342 accacgtccaccctgttcctgccttcagccaagtgtctctaccacgtccaccctgttc 401

QY 949 ctgcaactcctgacagcaacagatccttctgcaactctccccaggggtcacttgt 1008

```
Db      402 cctgcacacatggtgtgtgttgcattccaccacacattcccttcagccacggtct 461
Oy      1009 gtccagagacagcagcctagacactccttgagcagacagactcttccttcgacattc 1068
Db      462 accacgtccacccctgttccttcgacacagtggtgtgtgtgattccattccacacattc 521
Oy      1069 ccccaaggagctacttgtgtcaggagacagcagccttagcactccttcgagcaacag 1128
Db      522 ccttcagccacgcgtctctaccacgctccaccggttccttcgacacgltgggtgtgtaga 581
Oy      1129 ttc 1131
Db      582 tgc 584

RESULT 5
ABA64998
ID      ABA64998 standard; DNA; 611 BP.
XX
AC      ABA64998;
XX
DT      01-FEB-2002 (first entry)
XX
DE      Human foetal liver single exon nucleic acid probe #13303.
XX
KW      Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157277-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001MO-US000669.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-483447/52.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
PT      analyzing gene expression in human foetal liver -
XX
PS      Claim 4; SEQ ID NO 13303; 639pp + sequence listing; English.
XX
CC      The invention relates to a single exon nucleic acid probe for
CC      measuring human gene expression in a sample derived from human foetal
CC      liver. The single exon nucleic acid probes may be used for predicting,
CC      measuring and displaying gene expression in samples derived from human
CC      foetal liver. The present sequence is a single exon nucleic acid
CC      probe of the invention.
CC      Note: The sequence data for this patent did not form part of the
CC      printed specification, but was obtained in electronic format directly
CC      from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SO      Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;

Query Match      2.1%; Score 38.2; DB 22; Length 611;
Best Local Similarity 47.3%; Pred. No. 0.21;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Oy      889 aacagatccttcttgacacttcccccagagacacttgtgtcaggagacagcatatgc 948
```

```
Db      342 accacgtccacccctgttccttcgccccttcagccacgctgtctaccacgctcacctgttc 401
Oy      949 ctgacactctctgagacagacagatctcttccttcgacacttccccaggggtcacttgc 1008
Db      402 cctgcacacatggtgtgtgttgcattccaccacacattcccttcagccacggtct 461
Oy      1009 gtccagagacagcagccttagacactccttgagcagacagctccttcgacattc 1068
Db      462 accacgtccacccctgttccttcgacacagtggtgtgtgtgattccattccacacattc 521
Oy      1069 ccccaaggagctacttgtgtcaggagacagcagccttagcactccttcgagcaacag 1128
Db      522 ccttcagccacgcgtctctaccacgctccaccggttccttcgacacgltgggtgtgtaga 581
Oy      1129 ttc 1131
Db      582 tgc 584

RESULT 6
ABA32106
ID      ABA32106 standard; DNA; 611 BP.
XX
AC      ABA32106;
XX
DT      23-JAN-2002 (first entry)
XX
DE      Probe #10572 for gene expression analysis in human heart cell sample.
XX
KW      Human; gene expression; heart; microarray; vascular system; probe;
KW      cardiovascular disease; hypertension; cardiac arrhythmia;
KW      congenital heart disease; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157274-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001MO-US000666.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488899/53.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      hearts -
XX
PS      Claim 4; SEQ ID NO 10572; 530pp; English.
XX
CC      The present invention relates to single exon nucleic acid probes for
CC      measuring human gene expression in a sample derived from human heart. The
CC      present sequence is one such probe. The probes may be used for
CC      predicting, measuring and displaying gene expression in samples derived
CC      from the human heart via microarrays. By measuring gene expression, the
CC      probes are useful for predicting, diagnosing, grading, staging,
CC      monitoring and prognosing diseases of the human heart and vascular system
CC      e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC      congenital heart disease.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
```


XX WPI; 2001-488897/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID No 13854; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SNP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;
XX
Query Match 2.1%; Score 38.2; DB 22; Length 611;
Best Local Similarity 47.3%; Pred. No. 0.21;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
XX
OY 889 aacagatccttctgacactctcccccagagacacttgtgtcaggagacagcctgc 948
DB 342 aacacgtcacccctgttcctcctcagcagcagtgctctacacgtccaccctgttc 401
OY 949 ctgacactcctgacagcagacagatcctcttgacactctcccccagggtcacttgt 1008
DB 402 cctgcacacatgggtgtgtgtgtatccatccacacatcctcctcagcacagcgtct 461
OY 1009 gtacaggagacagcagcctagacactcctggtgagacagagtccttctgacacttc 1068
DB 462 accacgtcacccctgttcctcctcagcagcagtggtgtgtgtgtgtgtgtgtgtgt 521
OY 1069 ccccaagagatcacttgtgtcaggagacagcagcctagacactcctggtgagacag 1128
DB 522 cctcagcacagcagcgtctacacagctcacccggttccctgacacagtggtgtgtaga 581
OY 1129 tcc 1131
DB 582 tgc 584
XX
RESULT 11
AAI05680
ID AAI05680 standard; DNA; 611 BP.
XX
AC AAI05680;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #5671 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0633366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX
PS Claim 25; SEQ ID No 5671; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;
XX

Query Match 2.1%; Score 38.2; DB 22; Length 611;
Best Local Similarity 47.3%; Pred. No. 0.21;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
XX
OY 889 aacagatccttctgacactctcccccagagacacttgtgtcaggagacagcctgc 948
DB 342 accacgtcacccctgttcctcctcagcagcagtgctctacacagtcaccaccctgttc 401
OY 949 ctgacactcctgacagcagacagatcctcttgacactctcccccagggtcacttgt 1008
DB 402 cctgcacacatgggt 461
OY 1009 gtacaggagacagcagcctagacactcctggtgagacagagtccttctgacacttc 1068
DB 462 accacgtcacccctgttcctcctcagcagcagtggtgtgtgtgtgtgtgtgtgtgt 521
OY 1069 ccccaagagatcacttgtgtcaggagacagcagcctagacactcctggtgagacag 1128
DB 522 cctcagcacagcagcgtctacacagctcacccggttccctgacacagtggtgtgtaga 581
OY 1129 tcc 1131
DB 582 tgc 584
XX
RESULT 12
AAO03665
ID AAO03665 standard; DNA; 2465 BP.
XX
AC AAO03665;
XX
DT 07-AUG-1989 (first entry)
XX
DE Sequence homologous to Drosophila Per gene.
XX
KW Sex determination; ruminant embryos; ss.
XX
PN FR2635116-A.
XX
PD 09-FEB-1990.
XX
PF 08-AUG-1988; 88FR-0010706.
PR 08-AUG-1988; 88FR-0010706.
XX
PA (GEOR/) GEORGES M.
XX


```

XX OS Homo sapiens.
XX AC Key
XX FH CDS
XX FT 361..2740
XX FT /tag= a
XX FT /product=MAR subtype ml
XX FT misc-signal
XX FT 282
XX FT /tag= b
XX FT /label-splice acceptor site
XX FT /note=defines 5' end of exon
XX FT 2780
XX FT /tag= c
XX FT polyA-site
XX PN USN7241971-N.
XX PD 14-MAR-1989.
XX XX 08-SEP-1988; 88US-0241971.
XX PF 08-SEP-1988; 88US-0241971.
XX PR 08-SEP-1988; 88US-0241971.
XX XX (USSH ) NAT INST OF HEALTH.
XX PA WPI: 1989-165452/22.
XX DR P-PSDB; AAR06202.
XX DR
XX PT Cloned genes for muscarinic acetylcholine receptors -
XX PT for drug screening and diagnostic use.
XX PS Disclosure; ; P; English.
XX XX
XX CC The sequence was indexed from the best available specification copy;
XX CC the unidentifiable bases were annotated as 'N'.
XX CC The sequence may be useful for synthesis of hybridisation probes for
XX CC diagnostic use. The genes are cloned by screening a rat cerebral cortex
XX CC cDNA library with a probe based on nucleotides 170-225 of the non-coding
XX CC strand of porcine brain (ml) cDNA (modified at positions 5, 38 and 53);
XX CC identifying cDNA clones by hybridisation with BamHI or XhoI digests of
XX CC plasmid DNA from 12 cultures contg. up to 500000 independent clones;
XX CC recluturing until a pool contg. less than 5000 indep. clones with a
XX CC single hybridisation band are identified, and isolating individual
XX CC clones by colony hybridisation.
XX CC Stable cell lines produced by transfecting Chinese hamster ovary cells
XX CC (CHO-K1) with various pCD vectors contg. the gene inserts.
XX CC See also AAQ02068-Q020671 and AAQ00086.
XX CC (Note: Revised entry submitted to correct the patent number format of
XX CC US Government-owned NFIS applications to prevent clashes with ongoing US
XX CC granted patent numbers. For further information please visit the Derwent
XX CC web site at www.derwent.com/dwpi/updates/nfis_us.html.)
XX SQ Sequence 2880 BP; 618 A; 921 C; 763 G; 550 T; 28 other:

Query Match 2.0%; Score 37.2; DB 10; Length 2880;
Best Local Similarity 55.4%; Pred. No. 1.2;
Matches 72; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```

```

ID AAH99405 standard; cDNA; 1090 BP.
XX AAH99405;
XX AC
XX XX 16-OCT-2001 (first entry)
XX DT
XX XX
XX DE Human protein encoding cDNA sequence SEQ ID NO:240.
XX XX
XX KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX KW antibacterial; endocrine; cardiac; central nervous system; vitruide;
XX KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
XX KW antiagregant; haemostatic; vulnery; antilicer; osteopathic; eczema;
XX KW dermatological; antiallergic; antiasumatic; antibacterial; cytostatic;
XX KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
XX KW allergic rhinitis; diabetes; multiple sclerosis; depression;
XX KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX KW neurological disorder; ss.
XX OS
XX XX Homo sapiens.
XX AC WO200153455-A2.
XX PN
XX XX 26-JUL-2001.
XX PD
XX XX
XX PF 22-DEC-2000; 2000WO-US35017.
XX XX
XX PR 23-DEC-1999; 99US-0471275.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX XX
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI: 2001-457603/49.
XX DR P-PSDB; AAM25464.
XX DR
XX PT Isolated human polynucleotides encoding polypeptides, useful for the
XX PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX PS
XX XX Claim 1: Page 404-405; 1217pp; English.
XX XX
XX CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX CC AAM25963. The proteins can have activities based on the tissues and
XX CC cells they are expressed in, such as: antinflammatory; antirheumatic;
XX CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
XX CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;
XX CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;
XX CC antilicer; osteopathic; dermatological; antiallergic; antiasumatic;
XX CC antidabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX CC encoding them can be used in gene therapy, antisense therapy and vaccine
XX CC production. The proteins and polynucleotides are useful for screening for
XX CC agonists or antagonists of a protein and for the treatment and diagnosis
XX CC of disorders associated with the activity of a protein e.g. inflammation,
XX CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX CC osteoporosis, severe combined immunodeficiency, eczema, allergic
XX CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
XX CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
XX CC neurological disorders.
XX SQ Sequence 1090 BP; 296 A; 182 C; 223 G; 389 T; 0 other:

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 20:05:10 ; Search time 54.88 Seconds
(without alignments)
8190.767 Million cell updates/sec

Title: US-09-755-456-1

Perfect score: 1830
Sequence: 1 gaattcagatgcctcatacc.....tcccttagaagtgctgggattc 1830

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40.6	2.2	7218	1	US-08-232-463-14
2	34.8	1.9	3448	1	US-08-296-014A-3
3	34.8	1.9	3448	2	US-08-596-405-3
4	34.8	1.9	3448	1	US-08-877-620-3
5	34.8	1.9	4182	1	US-08-296-014A-1
6	34.8	1.9	4182	2	US-08-596-405-1
7	34.8	1.9	4182	2	US-08-877-620-1
8	33.8	1.8	1329	4	US-09-296-284-23
9	33.8	1.8	1337	4	US-09-296-284-2
10	33.8	1.8	4830	4	US-09-296-284-7
11	33.4	1.8	626	4	US-08-943-731-5
12	33.4	1.8	20084	4	US-08-943-731-199
13	33	1.8	1778	4	US-08-934-386-4
14	33	1.8	2661	4	US-08-592-126-70
15	33	1.8	5836	4	US-09-333-086-2
16	31.4	1.7	1672	2	US-09-003-199-1
17	31.2	1.7	1304	2	US-08-766-439-37
18	31.2	1.7	1304	2	US-08-766-439-38
19	31	1.7	1549	2	US-08-865-597A-1
20	30.8	1.7	1370	4	US-09-111-470-9
21	30.8	1.7	1458	4	US-09-111-470-3
22	30.8	1.7	7122	4	US-09-318-448-2
23	30.8	1.7	15297	4	US-09-817-180-3
24	30.6	1.7	1236	3	US-08-643-704A-48
25	30.6	1.7	2339	1	US-08-258-639A-1
26	30.6	1.7	2339	2	US-08-800-951-1
27	30.6	1.7	2339	5	PCT-US95-07391A-1

C 28	30.4	1.7	1212	2	US-09-092-770-18	Sequence 18, Appl
C 29	30.4	1.7	1212	4	US-09-222-851-18	Sequence 18, Appl
C 30	30.4	1.7	4858	3	US-08-436-332B-9	Sequence 9, Appl
C 31	30.2	1.7	4599	1	US-08-431-080-27	Sequence 27, Appl
C 32	30.2	1.7	4599	2	US-08-938-534-27	Sequence 27, Appl
C 33	30.2	1.7	5362	3	US-08-463-210-5	Sequence 5, Appl
C 34	30.2	1.7	12568	2	US-08-387-942C-1	Sequence 38, Appl
C 35	30	1.6	688	4	US-08-998-416-915	Sequence 915, App
C 36	30	1.6	2575	4	US-08-858-459A-11	Sequence 11, Appl
C 37	30	1.6	8220	2	US-08-568-459A-11	Sequence 11, Appl
C 38	30	1.6	8220	2	US-08-487-826B-13	Sequence 11, Appl
C 39	30	1.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 40	29.8	1.6	1137	3	US-09-082-088-1	Sequence 1, Appl
C 41	29.8	1.6	1137	4	US-09-546-117-1	Sequence 1, Appl
C 42	29.8	1.6	1368	4	US-09-200-673-13	Sequence 13, Appl
C 43	29.8	1.6	1501	1	US-08-349-025-1	Sequence 1, Appl
C 44	29.8	1.6	1501	2	US-08-566-096A-1	Sequence 1, Appl
C 45	29.8	1.6	1501	2	US-08-668-650B-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

```

Query Match 2.2%: Score 40.6; DB 1 Length 7218;
      Similarity 4.7%: Pred. No. 0.028;
Matches 19; Conservative 211; Mismatches 175; Indels 0; Gaps 0

OY 746 agtgcctccaccactctcgcctcagtcagtgactgtgagctgcacagctccatcgt 805
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1026 ATTATTTCCGACCTTGGCTGGCAGGTCGACGCGACGCTTGCATATTTTTTTTTT 1085

OY 806 tctgcagatgtctctcgtcgtctgctcaccaccagcagtcctcagctactacaatgg 865
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1086 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1145

OY 866 aaccacggtactctcgtcagcagcaagatctcttcgtcactctccccagggccac 925
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1146 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1205

OY 926 ttgtgtcagggacagggcctatgcactcctcgtcagcagcaagatctcttcgac 985
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1206 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1265

OY 986 tctcccccagggtcacttgttcaggagacagccagcctagcactcgtgagcaga 1045
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1266 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1325

OY 1046 acagtcctcttcgtcagctctccccaggatcacttgtgtcaggagacagccagcc 1105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1326 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1385

OY 1106 tagcactcctggcagcaacagtcctcttcgtcagccctccccc 1150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1386 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1450

RESULT 2
US-08-296-014A-3
: Sequence 3, Application US/08296014A
: Patent No. 5716834
:
: GENERAL INFORMATION:
: APPLICANT: Ding, Jeak Ling
: APPLICANT: Ho, Bow
: TITLE OF INVENTION: The Cloned Factor C cDNA of the
: TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscopus
: TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22042
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/296, 014A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy, Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 1781-105P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3448 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single

```

```

:      TOPOLOGY: both
:      MOLECULE TYPE: CDNA
:      HYPOTHETICAL: NO
:      ANTI-SENSE: NO
:      ORIGINAL SOURCE:
:      ORGANISM: Carcinoscorpius rotundicauda
:      FEATURE:
:      NAME/KEY: CDS
:      LOCATION: 18..3074
:      US-08-296-014A-3
:
:      Query Match      1.9%: Score 34.8; DB 1; Length 3448;
:      Best Local Similarity 65.4%; Fred.No.1.2;
:      Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0.
:
Oy  1197  gtctctgacacacagagacagctcagtgcacagtgctgtgcctccacagtcacacct 1256
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1078  gttctttgacacacctgtrtactgtgtggggtacagccatattaccatgacctttctcagctf 1137
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy  1257  gtcatgccgcacatcatg 1274
      ||||| ||||| |||||
Db  1138  gtcgtgacgacctcatg 1155
      ||||| ||||| |||||
:
RESULT      3
:      US-08-596-405-3
:      Sequence 3, Application US/08596405
:      Patent No. 5858706
:      GENERAL INFORMATION:
:      APPLICANT: Ding, Jeak Ling
:      APPLICANT: Ho, Bow
:      TITLE OF INVENTION: The Cloned Factor C CDNA of the
:      TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscorpius
:      TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
:      NUMBER OF SEQUENCES: 4
:      CORRESPONDENCE ADDRESS:
:      ADDRESSEE: Birch, Stewart, Kolasch & Birch
:      STREET: 8110 Gatehouse Road, Suite 500 East
:      CITY: Falls Church
:      STATE: Virginia
:      COUNTRY: USA
:      ZIP: 22042
:
:      COMPUTER READABLE FORM:
:      MEDIUM TYPE: floppy disk
:      COMPUTER: IBM PC compatible
:      OPERATING SYSTEM: PC-DOS/MS-DOS
:      SOFTWARE: PatentIn Release #1.0, Version #1.25
:      CURRENT APPLICATION DATA:
:      APPLICATION NUMBER: US/08/596,405
:      FILING DATE:
:      CLASSIFICATION: 435
:      ATTORNEY/AGENT INFORMATION:
:      NAME: Murphy, Jr., Gerald M.
:      REGISTRATION NUMBER: 28,977
:      REFERENCE/DOCKET NUMBER: 1781-105P
:      TELECOMMUNICATION INFORMATION:
:      TELEPHONE: (703) 205-8000
:      TELEFAX: (703) 205-8050
:      TELEX: 248345
:      INFORMATION FOR SEQ ID NO: 3:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 3448 base pairs
:      TYPE: nucleic acid
:      STRANDEDNESS: single
:      TOPOLOGY: both
:      MOLECULE TYPE: CDNA
:      HYPOTHETICAL: NO
:      ANTI-SENSE: NO
:      ORIGINAL SOURCE:
:      ORGANISM: Carcinoscorpius rotundicauda
:      FEATURE:
:      NAME/KEY: CDS

```


Db 1881 GTCGTCACGACCATCCATG 1898
|||||
RESULT 6
US-08-596-405-1
; Sequence 1, Application US/08596405
; Patent No. 5858706
; GENERAL INFORMATION:
; APPLICANT: Ding, Jeak Ling
; APPLICANT: HO, Bow
; TITLE OF INVENTION: The Cloned Factor C cDNA of the
; TITLE OF INVENTION: Singapore Horseshoe Crab, Carinoscorpius
; TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/596,405
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1781-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4182 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Carinoscorpius rotundicauda
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 569..3817
; US-08-596-405-1

Query Match 1.9%; Score 34.8; DB 2; Length 4182;
Best Local Similarity 65.4%; Pred. No. 1.4;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1197 gtctctgacagcaggaagctctagtgacagtgctgtgcccctagattccaaccct 1256
|||||
Db 1821 GTTCTTTGACAGCTGTACTGTGTGGGTACAGCCATATACCATGTAACCTTCTCAGTGT 1880

QY 1257 gtcatgcccgcacatcatg 1274
|||||
Db 1881 GTCGTCACGACCATCCATG 1898

;; GENERAL INFORMATION:
;; APPLICANT: Ding, Jeak Ling
;; APPLICANT: HO, Bow
;; TITLE OF INVENTION: The Cloned Factor C cDNA of the
;; TITLE OF INVENTION: Singapore Horseshoe Crab, Carinoscorpius
;; TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Birch, Stewart, Kolasch & Birch
;; STREET: 8110 Gatehouse Road, Suite 500 East
;; CITY: Falls Church
;; STATE: Virginia
;; COUNTRY: USA
;; ZIP: 22042
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/877,620
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/596,405
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murphy, Jr., Gerald M.
;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 1781-105P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 205-8000
;; TELEFAX: (703) 205-8050
;; TELEX: 248345
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4182 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: both
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Carinoscorpius rotundicauda
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 569..3817
;; US-08-877-620-1

Query Match 1.9%; Score 34.8; DB 2; Length 4182;
Best Local Similarity 65.4%; Pred. No. 1.4;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1197 gtctctgacagcaggaagctctagtgacagtgctgtgcccctagattccaaccct 1256
|||||
Db 1821 GTTCTTTGACAGCTGTACTGTGTGGGTACAGCCATATACCATGTAACCTTCTCAGTGT 1880

QY 1257 gtcatgcccgcacatcatg 1274
|||||
Db 1881 GTCGTCACGACCATCCATG 1898

RESULT 8
US-09-296-284-23
; Sequence 23, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes

```

? TITLE OF INVENTION: and Methods of Use Thereof
?
? FILE REFERENCE: 1533.0870000
?
? CURRENT APPLICATION NUMBER: US/09/296,284A
?
? CURRENT FILING DATE: 1999-04-22
?
? NUMBER OF SEQ ID NOS: 87
?
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 23
?
? LENGTH: 1329
?
? TYPE: DNA
?
? ORGANISM: Gluconobacter suboxydans
?
? US-09-296-284-23

```

Query Match	1.88;	Score 33.8;	DB 4;	Length 1329;
Best Local Similarity	50.38;	Pred. NO. 1.4;		
Matches 83; Conservative	0;	Mismatches 82;	Indels 0;	Gaps 0;

Accession	Sequence	Position
Oy	1150 cagggagccactctgtgcaggaaccaagatctcggaaacacggagctctctcgagacg	1209
Db	1140 ctggggcaacccgagctccggagaccctctgctcccgatctcggaaagctccgacatc	1199
Oy	1210 agggagcagctctagtgagacgctgtgtgccccccagatccaaacccctgtcatgcccgcac	1269
Db	1200 gggctctcgcggttcccaacggccggcgtggaagctctcttccaaaggcctgtgagttccatc	1259
Oy	1270 ccataaacattagtggaatatggtatccctctcttaccgcgacata	1314
Db	1260 gccgagacctttagcgaaagctctgacatctcccccgcagaaacata	1304

RESULT 9
US-09-296-284-2
; Sequence 2, Application US/09296284A
; Date of Invention 20090909

; GENERAL INFORMATION:

APPLICANT: Chol, Eun-Sung

APPLICANT: Lee, Eun-Hae

and Methods of Use Thereof

CURRENT APPLICATION NUMBER: US/09/296,284A

NUMBER OF SEQ ID NOS: 87

```

; SOFTWARE: PatentIn Ver. 2.0.0

```

; LENGTH: 1437

TYPE: DNA

ORGANISM: GLUCONOBACTER SUBOXIDANS
DS-09-296-284-2

Query Match	1.88;	Score 33.8;	DB 4;	length 1437;
Best Local Similarity	50.38;	Pred. No. 1.5;		
Matches	83;	Conservative	0;	Mismatches 82;
			Indels	0;
			Gaps	0

QY	1150	QY	1248	QY	1210	QY	1308	QY	1270	QY	1368
caaggagacgaacttctgtgcacaggaacaaagatgctgccttgaaacatccggagatctcttcgacagc	1209	ctggcgacacacgaagctctccgggaacccctgtctgcctcgtgatactcgcaagctccgcacatc	1307	agggagacagctctcagtgagacaaatgctgtgcgccctcagaattcttcaaacacctgtgcacgcgcgcat	1265	gggtactcgcggcttccacagcgcgcgcgtggaaagctctctcttccaaagggctctgaaatgacctaatc	1367	ccatgacacataatgtagaanaatgggtatctctctcttctactacgcgcataca	1314	gcgcgaagccttatgctcgaaagcctgcggacactctctctcccgacagaaaca	1412

RESULT 10
US-09-296-284-7
; Sequence 7, Application US/09296284A
; Patent No. 6204040

```

: GENERAL INFORMATION:
: APPLICANT: Choi, Eun-Sung
: APPLICANT: Rhee, Sang-Ki
: APPLICANT: Lee, Eun-Hae
: TITLE OF INVENTION: Glucuronobacter Suboxydans Sorbitol Dehydrogenase, Genes
: TITLE OF INVENTION: and Methods of Use Thereof
: FILE REFERENCE: 1533.0870000
: CURRENT APPLICATION NUMBER: US/09/296,284A
: CURRENT FILING DATE: 1999-04-22
: NUMBER OF SEQ. ID NOS: 87
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 7
: LENGTH: 4830
: TYPE: DNA
: ORGANISM: Glucuronobacter suboxydans
US-09-296-284-7

```

Query Match	1.88;	Score 33.8;	DB 4;	Length 4830;
Best Local Similarity	50.38;	Pred. No. 3.1;		
Matches 83;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;

QY	1150	caggagacgacctgtgtgcagaagatcgaatgctctgaaacacgcgaatctctgcagacgc	1209
Db	4211	ctggggaacacagagctccgggaacctgtctgtcctcgatctcgcaagctccgacatc	4270
QY	1210	agggaacgctctagtgaggacatgtgtgcccctcaagatctcaaacctgtcatgctgcgcac	1269
Db	4271	gggtacctcgcggtctccaccggcgcgcgtggaagctctctctccaaaggctcgtgaagtgctctacat	4330
QY	1270	ccatgaacattagtgaagaatggatctctctctctctactacgcgcataca	1314
Db	4331	gcgcgagcctctatgycgaagctcgcgaacctctctcccgacgaacaca	4375

RESULT 11
US-08-943-731-199/c
; Sequence 199, Application US/08943731

Patent NO. 6265157

APPLICANT: PROCKOP, DARWIN J.

APPLICANT: DELTAS, CONSTANTINOS D.

APPLICANT: LARSON, ANDREA W.

APPLICANT: COLIGE, ALAIN

```

; AFFILIATION:
; NAME: JAMES
; APPLICANT: KORRKO, JARMO

```

APPLICANT: ALA-NOKKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING

NUMBER OF SEQUENCES:	666
TITLE OF INVENTION:	ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES

CORRESPONDENCE ADDRESS: PANT

STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FIP

CITY: PHILADELPHIA
STATE: PA

COUNTRY: USA
STD: 10103-7

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/943,731

CLASSIFICATION: 435

APPLICATION NUMBER: US 08/212,322

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 199:
SEQUENCE CHARACTERISTICS:
LENGTH: 626 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-199

Query Match 1.8%; Score 33.4; DB 4; Length 626;
Best Local Similarity 57.0%; Pred No. 1.2;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1107 agcactcctgggcaagaaggtccttcttggcaccccttcccaaggagcacttgtgt 1166
||| | ||||| |||| | ||| ||| ||| |||
Db 537 ACCTATGTTGGCGAGCGGCGACCCCGACCAACACCCCTGGGCTGTGGGTGC 478

Qy 1167 caggatcaaggtatgcttggaaatccgcgaggttctctggacagagg 1213
||| | ||| ||| |||| | ||| ||| ||| |||
Db 477 GTGAGAAACTGTGTGTGATCCACAGATGCTGCAGAGGACGCACGG 431

RESULT 12
US-08-943-731-5/C
Sequence 5, Application US/08943731
Patent No. 6285157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISSA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR.
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 20084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-5

Query Match 1.8%; Score 33.4; DB 4; Length 20084;
Best Local Similarity 57.0%; Pred No. 10;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1107 agcactcctgggcaagaaggtccttcttggcaccccttcccaaggagcacttgtgt 1166
||| | ||||| |||| | ||| ||| ||| |||
Db 13825 ACCTATGTTGGCGAGCGGCGACCCCGACCAACACCCCTGGGCTGTGGGTGC 13766

Qy 1167 caggatcaaggtatgcttggaaatccgcgaggttctctggacagagg 1213
||| | ||| ||| |||| | ||| ||| ||| |||
Db 13765 GTGAGAAACTGTGTGTGATCCACAGATGCTGCAGAGGACGCACGG 13719

RESULT 13
US-08-934-386-4/C
Sequence 4, Application US/08934386
Patent No. 6306636
GENERAL INFORMATION:
APPLICANT: Haselkorn, Robert
APPLICANT: Gornicki, Piotr
TITLE OF INVENTION: Methods for Detecting Nucleic Acid
TITLE OF INVENTION: Segments Encoding Acetyl-CoA Carboxylase
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 433
CITY: Houston
STATE: Texas
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,386
FILING DATE: 19-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARSB:521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1778 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-934-386-4

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 21:33:31 ; Search time 1982.69 Seconds
(without alignments)
12457.534 Million cell updates/sec

Title: US-09-755-456-1

Perfect score: 1830

Sequence: 1 gaattcagatgctcatcac.....tcccttagaagtgtgagatc 1830

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthum:*
3: em_estlin:*
4: em_estinu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	2.3	720	12	BH535822 BOGEY58TF
2	42.6	2.3	685	12	BH434901 BOHNO30TF
3	41.2	2.3	387	10	BM280783 K105601.Y
4	40	2.2	1000	12	CNS000COQ
5	39.2	2.1	498	10	BM281363
6	39.2	2.1	547	9	AW165709 MBASBMA02
7	39.2	2.1	565	10	BM280570 K102610.Y
8	39.2	2.1	566	10	BM280471 K101510.Y
9	39.2	2.1	577	10	BM285287 kh96h09.Y
10	39.2	2.1	614	10	BM284986
11	39.2	2.1	633	10	BM280716 K104D10.Y
12	39.2	2.1	644	9	AW165781 MBASBMA19
13	39	2.1	549	10	BI282815 UI-R-DA0-
14	39	2.1	699	12	CNS020KA
15	38.6	2.1	311	10	BG955899
16	38.4	2.1	1034	12	CNS03MEV
17	38.2	2.1	405	10	H16015 Y127C03.r1

C 18	38.2	2.1	853	12	CNS03VR2	AL262199 Tetradon
C 19	38	2.1	489	10	BF773988	BF773988 283520 MA
C 20	37.6	2.1	757	12	CNS02J08	AL200537 Tetradon
C 21	37.6	2.1	803	12	CNS026N7	AL183580
C 22	37.4	2.0	545	12	AQ432826	AQ432826 HS-5132-B
C 23	37.4	2.0	626	10	BI279337	BI279337 UI-R-DA0-
C 24	37	2.0	326	9	AV936678	AV936678
C 25	37	2.0	421	10	BI397496	BI397496 AC95 AC-A
C 26	37	2.0	1080	12	CNS05IBJ	AL338680 Tetradon
C 27	36.8	2.0	726	10	BF009183	BF009183 601873016
C 28	36.8	2.0	781	10	BG564009	BG564009 602584842
C 29	36.8	2.0	880	10	BE973122	BE973122 601652003
C 30	36.6	2.0	513	10	BG383812	BG383812 302198 MA
C 31	36.6	2.0	849	12	CNS011T7	AL100693 Drosophila
C 32	36.4	2.0	584	9	AJ392991	AJ392991
C 33	36.4	2.0	899	9	AL537015	AL537015
C 34	36.4	2.0	986	9	AL533429	AL533429
C 35	36.2	2.0	473	12	AQ116963	AQ116963 HS-3021-A
C 36	36.2	2.0	723	12	CNS01UY4	AL168241 Tetradon
C 37	36.2	2.0	802	12	CNS03YUV	AL262912 Tetradon
C 38	36.2	2.0	1101	12	CNS00L7A	AL068159 Drosophila
C 39	36	2.0	343	9	AI081631	AI081631 ow/77603.s
C 40	36	2.0	677	10	BF703878	BF703878 MI-P-01-a
C 41	36	2.0	685	9	AU132863	AU132863
C 42	36	2.0	965	12	CNS01ZDK	AL174161 Tetradon
C 43	36	2.0	1056	10	BG515679	AL214842 Tetradon
C 44	35.8	2.0	466	10	BG515679	BG515679 de03d07.
C 45	35.8	2.0	469	10	BG513854	BG513854 de03d07.

ALIGNMENTS

RESULT 1
BH535822
LOCUS BH535822 720 bp DNA linear GSS 14-DEC-2001
DEFINITION BOGEY58TF BOGE Brassica oleracea genomic clone BOGEY58, DNA
sequence.
ACCESSION BH535822
VERSION BH535822.1 GI:17767517
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 720)
TOWN, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other-GSSs: BOGEY58TF
COMMENT Contact: Chris Town
TIGR
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
location/Qualifiers
1..720
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOGEY58"
/clone_11b="BOGE"
/note="Vector: pGEM1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pGEM1 using BstXI linkers"

FEATURES

BASE COUNT
263 a 100 c 101 g 256 t
ORIGIN

	Query Match	Similarity	2.3%	Score	43	DB	12	Length	720
	Best Local	Similarity	53.2%	Pred.	No.	0.4%			
	Matches	91	Conservative	0	Mismatches	80	Indels	0	Gaps
Oy	220	tttcaacctcgaatggaacagagatgaactctctgtggaataatcgtatagatc	279						
Db	421	TTTTTAACATATAATTTTAAACAAAGATTAATTTATTATTATTAATTAATTAACAAAGTCMAA	480						
Oy	280	ggagacattggaaccacggagattgaaataaagtataaattccctccgaaaacgcaagt	339						
Db	481	CATTAACATTTTAAACATATGATTTTAAATTTAAATAAACAATGAACATTAATAAACAATAATGTTG	540						
Oy	340	ttaagatgaagaagtatttcttaagatccctctctgcctctatgcccctgatga	390						
Db	541	TAAATTAACATTAATTTTCAAGAGAGCATCGAAGCTCAGTTTGTGCTTTGA	591						

RESULT	2
BH434901/c	
LOCUS	BH434901 685 bp DNA linear GSS 12-DEC-2001
DEFINITION	BOHNQ30TF BOHN Brassica oleracea genomic clone BohnQ30, DNA sequence.
ACCESSION	BH434901
VERSION	BH434901
KEYWORDS	BH434901.1 GI:17620622
SOURCE	GSS.
ORGANISM	Brassica oleracea. Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 685) Town,C.D., Van Aken,S., Uterback,T. and Fraser,C.M. Whole genome shotgun sequencing of Brassica oleracea Unpublished (2001) Other_GSSs: BOHNQ30TR Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TF Class: sheared ends.
FEATURES	location/qualifiers
source	1..685 /organism="Brassica oleracea" /strain="TO1000DH3" /db_xref="taxon:3712" /clone="BOHNQ30" /clone_lib="BOHN" /note="vector: pROS1; site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pROS1 using BstXI linkers"
BASE COUNT	189 a 105 c 144 g 247 t
ORIGIN	

Query Match	2.3%	Score 42.6	DB 12	Length 685	
Best Local Similarity	51.9%	Pred. NO. 0.62			
Matches	96	Conservative	0	Mismatches 89	
				Indels 0	
				Gaps 0	
QY	239	aacagagatgaac	ctctgtgcaaatcg	tagatcgcattgagacattgaacacg	298
Db	684	ATCATTCAATGAAC	CTTTCTATGAAG	ATATTCACACCAAAATCAAACTTCCCAAGTACT	625
QY	239	agttgaaataa	gataaatacctccg	aaacgcagagatttaagatgaagaagtatttc	358
Db	624	ACATTATCATATAAT	TGGCACAATATGCCATTAACAGAGAGTAAAGTTCTCGATTCCTTG	565	
QY	359	taagatcctctctg	ctctattatccctcg	ataaagttagaagtgcctattcaagaacgc	418
Db	564	CATATCCTTCACAA	TGTCGAGGTTCC	CACACGAGAAAGTTACCAATGTGCTCCGATATTCACC	505

QY	419	actgt	423
Db	504	AGTGT	500

RESULT	3
LOCUS	BM280783
DEFINITION	BM280783 387 bp mRNA linear EST 28-DEC-2001 k105a01.v1 Ascaris suum i4 pSOPRI1 zarlenga v1 Ascaris suum cDNA 5 similar to wp:CE25434 Y54E10B1.2 collagen , mRNA sequence.
ACCESSION	BM280783
VERSION	BM280783.1 GI:17989825
KEYWORDS	EST.
SOURCE	p1g roundworm.
ORGANISM	Ascaris suum Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
REFERENCE	1 (bases 1 to 387) McCartter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,T., Wylie,T. Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y. Gibbons,M., Rittler,E., Bennett,J., Franklin,C., Tsegatsivili,R. Ronko,I., Kennedy,S., Maguire,K., Beck,C., Underwood,K., Steptoe .M., Allen,M., Pearson,B., Swaller,T., Harvey,N., Schurr,K., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE	The Washington Univ. Nematode EST Project, 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: McCartter JP

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was supplied by Dr. Dante Zarlenga of the USA
Immunology and Disease Resistance lab in Beltsville, MD. The cDNAs
were made from 21 day 14 proximal (isolated from jejunum) and
directionally cloned into the NotI/SalI sites of pSforT1.
Seq primer: -40RP from Gibco
High quality sequence stop: 386.

FEATURES	SOURCE	LOCATION/Qualifiers
		1. .387
		/organism="Ascaris suum"
		/db_xref="taxon:6253"
		/clone_lib="Ascaris suum L4 pSPOR1 Zarlenga v1"
		/dev_stage="L4"
		/lab_host="DH5-alpha"
		/note="Vector: pSPOR1 (Life Technologies); Site_1: NotI;
		Site_2: SalI; The library was supplied by Dr. Dante
		Zarlenga of the USDA Immunology and Disease Resistance Lab
		in Beltsville, MD. The cDNAs were made from 21 day L4
		proximal (isolated from jejunum) and directionally cloned
		into the NotI/SalI sites of pSPOR1."
BASE COUNT		89 a 125 c 110 g 63 t
ORIGIN		

Query Match	2.3%	Score 41.2;	DB 10;	Length 387;
Best Local Similarity	50.5%;	Pred. No. 1.3;		
Matches 100;	Conservative	0;	Mismatches 98;	Indels 0;
			Gaps	0;
QY 1048	aggtccttccttgcgcactttcccccaggagtaactttgtctcagggagacagggccacctta	1107		
Db 141	ATGGACCTCCCGGTGCACACGCAAAAGATGGTGTCTCCCGTGTATGAGACTTGCCGCCAG	200		
QY 1108	gcactcctctgggcagcaacaggatcctttcttcttgcgcaccttctcccaaggagacattgtgtc	1167		
Db 201	GACCTCCGAGGACAGAGATGCCGAACTTTCACAGATGCCCTTCTGGCCCTTACACACAAATGTC	260		
QY 1168	aggtatcaagatgatgcctcttgaacatcccgagatttctctgagcagcagaggaacatctagtggac	1227		
Db 261	CGTGCAAGCTCTCTCTGGACCTCTCTGGGCGCTCTCTGGAGCACACTGGACAGATGAGTGCACAC	320		

Db 403 CCTGTCAACCTCCTCCTGGACCTCTGGGCTCTCTGACCACTGGACAGATGACAAC 462

RESULT 6
 AM165709
 LOCUS
 DEFINITION MBASBM022M13R Ascaris suum (parasitic nematode) body wall muscle and hypodermis Ascaris suum cDNA clone MBASBM022 5', mRNA sequence.
 ACCESSION AM165709
 VERSION AM165709.1 GI:6382584
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.
 1 (bases 1 to 547)
 Damb,J., Geary,T. and Blaxter,M.
 A survey of genes expressed in the parasitic nematode Ascaris suum
 Unpublished (2000)
 JOURNAL
 COMMENT
 TITLE
 INSTITUTE OF CELL, ANIMAL AND POPULATION BIOLOGY
 CONTACT: BLAXTER ML
 UNIVERSITY OF EDINBURGH
 ASHWORTH LABS, KING'S BUILDINGS, WEST MAINS ROAD, EDINBURGH, EH9
 3JT, UK.
 TEL: +44 131 650 6760
 FAX: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The Ascaris suum EST dataset (including the ASC clustering
 information) is available on the www at http://www.ed.ac.uk/tlide
 (mbx/Ascarisweb/AscarisEST.html)
 PCR Primers
 FORWARD: M13 Reverse (AGCGGATACCAATTTCACACAGCA)
 BACKWARD: M13 Forward (CGCGGATTTTCCACGACAGC)
 Seq primer: M13 Reverse (AGCGGATACCAATTTCACACAGCA).

FEATURES

Source

Location/Qualifiers

1..547
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone_lib="MBASBM022"
 /clone_1ib="Ascaris suum (parasitic nematode) body wall
 muscle and hypodermis"
 /sex="mixed"
 /tissue_type="body wall muscle and hypodermis"
 /dev_stage="adult"
 /note="Vector: Lambda Zap II; Site_1: NotI (5'end);
 Site_2: NotI (3'end); Ascaris suum is an intestinal
 nematode parasite of pigs. The library was constructed
 from dissected body wall muscle and hypodermis tissue for
 Dr. T. Geary, Pharmacia-Upjohn Inc, Kalamazoo, MI, USA
 (tgeary@am.pnu.com)"

BASE COUNT

121 a 172 c 152 g 102 t

ORIGIN

Query Match 2.1%; Score 39.2; DB 9; Length 547;
 Best Local Similarity 51.1%; Pred. No. 5.6;
 Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtccttctggaactcttccccagagatcatttggtaagagagcgacgctta 1107
 DB 242 ATGGACCTCCGGGTGACCAAGCAAGATGATGTCCTCCGCTGATGACCTGGCCCGCAG 301
 QY 1108 gaactcctggagcaacaagtccttctggaaccttcccaagagcacttggta 1167
 DB 302 GACCTCCAGACCAAGATGCGCAATTCACAGATCGTCTTGTCCCGCACCAACACATGTC 361
 QY 1168 aggatcaagatgatgcctgtaacatccgaggttcctggaagagagagcagtcagtgcac 1227
 DB 362 CCTGTCAACCTCCTCCTGGACCTCTGGGCTCTCTGACCACTGGACAGATGACAAC 421

RESULT 7

BM280570
 LOCUS
 DEFINITION BM280570 565 bp mRNA linear EST 28-DEC-2001
 K102D10.y1 Ascaris suum L4 pSPORT1 Zarlenga v1 Ascaris suum cDNA 5'
 similar to WP:CE25434 Y54E10BL.2 collagen ;, mRNA sequence.
 ACCESSION BM280570
 VERSION BM280570.1 GI:17989612
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.
 1 (bases 1 to 565)
 McCarter,D., Clifton,S., Chapell,B., Pape,D., Martin,J., Wyle,T.,
 Dante,M., Maira,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
 Gibbons,M., Rutter,E., Bennett,J., Franklin,C., Tsagarelisvill,R.,
 Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
 ,M., Allen,M., Pearson,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
 Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
 Wilson,R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 JOURNAL
 COMMENT
 TITLE
 CONTACT: MCCARTER JP
 THE WASHINGTON UNIV. NEMATODE EST PROJECT, 1999
 WASHINGTON UNIVERSITY SCHOOL OF MEDICINE
 4444 FOREST PARK PARKWAY, BOX 8501, ST. LOUIS, MO 63108, USA
 TEL: 314 286 1800
 FAX: 314 286 1810
 Email: est@watson.wustl.edu
 The library was supplied by Dr. Dante Zarlenga of the USDA
 Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs
 were made from 21 day L4 proximal (isolated from jejunum) and
 directionally cloned into the NotI/SalI sites of pSPORT1.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 481.

FEATURES

Source

Location/Qualifiers

1..565
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone_1ib="Ascaris suum L4 pSPORT1 Zarlenga v1"
 /dev_stage="L4"
 /lab_host="DH5-alpha"
 /note="Vector: pSPORT1 (Life Technologies); Site_1: NotI;
 Site_2: SalI; The library was supplied by Dr. Dante
 Zarlenga of the USDA Immunology and Disease Resistance Lab
 in Beltsville, MD. The cDNAs were made from 21 day L4
 proximal (isolated from jejunum) and directionally cloned
 into the NotI/SalI sites of pSPORT1."

BASE COUNT

118 a 176 c 157 g 114 t

ORIGIN

Query Match 2.1%; Score 39.2; DB 10; Length 565;
 Best Local Similarity 51.1%; Pred. No. 5.7;
 Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtccttctggaactcttccccagagatcatttggtaagagagcgacgctta 1107
 DB 270 ATGGACCTCCGGGTGACCAAGCAAGATGATGTCCTCCGCTGATGACCTGGCCCGCAG 329
 QY 1108 gaactcctggagcaacaagtccttctggaaccttcccaagagcacttggta 1167
 DB 330 GACCTCCAGACCAAGATGCGCAATTCACAGATCGTCTTGTCCCGCACCAACACATGTC 389
 QY 1168 aggatcaagatgatgcctgtaacatccgaggttcctggaagagagagcagtcagtgcac 1227
 DB 390 CCTGTCAACCTCCTCCTGGACCTCTGGGCTCTCTGACCACTGGACAGATGACAAC 449

RESULT 8
 BM280471
 LOCUS
 DEFINITION BM280471 566 bp mRNA linear EST 28-DEC-2001
 K101B10.y1 Ascaris suum L4 pSPORT1 Zarlenga v1 Ascaris suum cDNA 5'
 similar to WP:CE25434 Y54E10BL.2 collagen ;, mRNA sequence.

ACCESSION	BM280471	
VERSION	BM280471.1	GI:17989506
KEYWORDS	EST.	
SOURCE	p19 roundworm.	
ORGANISM	Ascaris suum	
	Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea	
REFERENCE	1 (bases 1 to 566)	
AUTHORS	McCarter,J., Clifton,S., Chlapell,B., Page,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisvilli,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,R., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., Mcann,R., Waterston,R. and Wilson,R.	
TITLE	The Washington Univ. Nematode EST Project, 1999	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: McCarter JP The Washington Univ. Nematode EST Project, 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1. Seq primer: -40RP from Glbco High quality sequence stop: 480. Location/Qualifiers 1..566 /organism="Ascaris suum" /db_xref="taxon:6253" /clone_lib="Ascaris suum L4 pSPORT1 zarlenga v1" /dev_stage="L4" /lab_host="DH5-alpha" /note="Vector: pSPORT1 (Life Technologies); Site_1: NotI; Site_2: SalI; The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1."	
FEATURES		
Source		
BASE COUNT	119 a 174 c 158 g 115 t	
ORIGIN		
Query Match	2.1%; Score 39.2; DB 10; Length 566;	
Best Local Similarity	51.1%; Pred. No. 5.7;	
Matches	92: Conservative 0; Mismatches 88; Indels 0; Gaps 0;	
OY	1048 aggtccttcctgagcactctcccgagagtcacttgatcagagagcgacgacctta 1107	
Db	271 ATGAGACCTCCGGGTGACACGAGCAAAAGTGGTGTCCCGGTATGAGCTGCGCCGACG 330	
OY	1108 gcactcctcctggcagcaacaggtctcttcttcctgagcactctcccgaggagccattgtgtc 1167	
Db	331 GACCTCCAGGAGACGACATGCCGAACCTTCAAGATCGCTCTGCGCCGTACACCAACATGTC 390	
OY	1168 aggaatcaaggtatgctctggaacatccggaagtctctgagcagcagaggaagcttagtgagc 1227	
Db	391 CCTGTCAAGCTCTCTCTGAGACTCTCTGTGGCGCTCCCTGTGACACCACTGTGAGATGACAAC 450	
RESULT	9	
LOCUS	BM285287	
DEFINITION	kh56h09.v1 Ascaris suum L4 pSPORT1 zarlenga v1 Ascaris suum cDNA 5' rML	
ACCESSION	BM285287	
VERSION	BM285287.1	GI:17994329
KEYWORDS	EST.	

FEATURES	source
BASE COUNT	122 a 182 c 156 g 117 t
ORIGIN	
Query Match	2.1%; Score 39.2; DB 10; Length 577;
Best Local Similarity	51.1%; Pred. NO. 5.7;
Matches	92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
OY 1048	aggtccttcctggaacctctccccaagagtcacttgctgtaaggaagcaccagccta 1107
DB 289	ATGGACCTCCGGGTGACACAGCAAGATGGTGTCCCGTATGATGACCGCCGCGAC 348
OY 1108	gcacccctgggagagaacagcgtctcttcgtgacacctctcccaaggagccacttggtc 1167
DB 349	GACCTCCAGAGCAGATGCCGACATTCACAGATCGTCTCTGCCCCCTACCAACATGTC 408
OY 1168	aggaatcaagatgactgactgaaacatccggaggtctcggaagcaggaagcagtcagtgcac 1227
DB 409	CCTGTCAAGCTCCTCCTGTGACCTCCTGTGGCCCTCCGTGAGACCACTGTGACAGATGACAC 468
RESULT 10	
LOCUS	BM284986 614 bp mRNA linear EST 28-DEC-2001
DEFINITION	kb03d05.y1 Ascaris suum L4 psf00r1 Zarlenga v1 Ascaris suum cDNA, 5' similar to wp:CE25434 Y54E10B1.2 collagen; contains MER22.b1 MSRI repetitive element; mRNA sequence.
ACCESSION	BM284986
VERSION	BM284986.1 GI:17994028
KEYWORDS	EST.
SOURCE	plg roundworm.
ORGANISM	Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea
TITLE	The Washington Univ. Nematode EST Project, 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: McCarter JP The Washington Univ. Nematode EST Project, 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPOR1. Seq primer: -40RP from Glbco High quality sequence stop: 499. Location/Qualifiers 1..577 /organism="Ascaris suum" /db_xref="taxon:6253" /clone_11b="Ascaris suum L4 psf00r1 Zarlenga v1" /dev_stage="L4" /lab_host="DH5-alpha" /note="Vector: pSPOR1 (Life Technologies); Site_1: NotI; Site_2: SalI. The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPOR1."

REFERENCE
AUTHORS
1 (bases 1 to 614)
McCartier,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
Dante,M., Mairs,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisvill,R.,
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCartier JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was supplied by Dr. Dante Zarlenga of the USDA
Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs
were made from 21 day I4 proximal (isolated from jejunum) and
directionally cloned into the NotI/Sali sites of pSPORT1.
Seq primer: -40RP from Gibco
High quality sequence stop: 483.
Location/Qualifiers
1. .614
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_lib="Ascaris suum I4 pSPORT1 Zarlenga v1"
/dev_stage="I4"
/lab_host="DH5-alpha"
/note="Vector: pSPORT1 (Life Technologies); Site_1: NotI;
Site_2: Sali; The library was supplied by Dr. Dante
Zarlenga of the USDA Immunology and Disease Resistance Lab
in Beltsville, MD. The cDNAs were made from 21 day I4
proximal (isolated from jejunum) and directionally cloned
into the NotI/Sali sites of pSPORT1."

BASE COUNT
ORIGIN
129 a 196 c 170 g 119 t

Query Match
Best Local Similarity 51.1%; Score 39.2; DB 10; Length 614;
Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtccttcttgacactcttccccagagatcacttctgtgaagagacagccacta 1107
DB 268 ATGAGCCTCCGGGTGCACAGCAAGATGATGCTCCGGTATGACCTGGCCGCCAG 327
QY 1108 gaactcctggagcagaacagtccttcttgacaccttctccagagaccacttctgtc 1167
DB 328 GACCTCCAGGACCAATGCCGATTCACAGATCGTCTTCTGCCCCGACCAACAATGTC 387
QY 1168 aggatcaagatagctctggaacatccgagatcctcctggagacagagacagctctagtgac 1227
DB 388 CCTGTCAAGCTCCTCTGGACCTCCTGGGCTCTCTGGACCACTGGAGATGAGCAAC 447

RESULT 11
BM280716 633 bp mRNA linear EST 28-DEC-2001
LOCUS K104b10.v1 Ascaris suum I4 pSPORT1 Zarlenga v1 Ascaris suum cDNA 5'
DEFINITION similar to WP:CE25434 Y54E10B.2 collagen 1; Ascaris suum sequence.
ACCESSION BM280716
VERSION BM280716.1 GI:17989758
KEYWORDS EST.
SOURCE pig roundworm.
ORGANISM Ascaris suum
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea
; Ascarididae; Ascaris.
REFERENCE
AUTHORS McCartier,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
Dante,M., Mairs,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisvill,R.,
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.

Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisvill,R.,
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCartier JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was supplied by Dr. Dante Zarlenga of the USDA
Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs
were made from 21 day I4 proximal (isolated from jejunum) and
directionally cloned into the NotI/Sali sites of pSPORT1.
Seq primer: -40RP from Gibco
High quality sequence stop: 501.
Location/Qualifiers
1. .633
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_lib="Ascaris suum I4 pSPORT1 Zarlenga v1"
/dev_stage="I4"
/lab_host="DH5-alpha"
/note="Vector: pSPORT1 (Life Technologies); Site_1: NotI;
Site_2: Sali; The library was supplied by Dr. Dante
Zarlenga of the USDA Immunology and Disease Resistance Lab
in Beltsville, MD. The cDNAs were made from 21 day I4
proximal (isolated from jejunum) and directionally cloned
into the NotI/Sali sites of pSPORT1."

BASE COUNT
ORIGIN
132 a 199 c 173 g 129 t

Query Match
Best Local Similarity 51.1%; Score 39.2; DB 10; Length 633;
Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtccttcttgacactcttccccagagatcacttctgtgaagagacagccacta 1107
DB 289 ATGAGCCTCCGGGTGCACAGCAAGATGATGCTCCGGTATGACCTGGCCGCCAG 348
QY 1108 gaactcctggagcagaacagtccttcttgacaccttctccagagaccacttctgtc 1167
DB 349 GACCTCCAGGACCAATGCCGATTCACAGATCGTCTTCTGCCCCGACCAACAATGTC 408
QY 1168 aggatcaagatagctctggaacatccgagatcctcctggagacagagacagctctagtgac 1227
DB 409 CCTGTCAAGCTCCTCTGGACCTCCTGGGCTCTCTGGACCACTGGAGATGAGCAAC 468

RESULT 12
AM165781 644 bp mRNA linear EST 12-NOV-1999
LOCUS MBASBM197M13R Ascaris suum (parasitic nematode) body wall muscle
DEFINITION and hypodermis Ascaris suum cDNA clone MBASBM197 5', mRNA
sequence.
ACCESSION AM165781
VERSION AM165781.1 GI:6382712
KEYWORDS EST.
SOURCE pig roundworm.
ORGANISM Ascaris suum
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea
; Ascarididae; Ascaris.
REFERENCE
AUTHORS Daub,J., Geary,T., and Blaxter,M.
TITLE A survey of genes expressed in the parasitic nematode Ascaris suum
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology

University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The Ascaris suum EST dataset (including the ASC clustering
information) is available on the www at <http://www.ed.ac.uk/tilde/mjb/Ascarisweb/AscarisEST.html>
PCR Primers
FORWARD: M13 Reverse (AGCGATTAACATTTCACACAGA)
BACKWARD: M13 Forward (CGCCAGGTTTCCGATCAGAC)
Seq primer: M13 Reverse (AGCGATTAACATTTCACACAGA).
Location/Qualifiers
1. 644

FEATURES

source

/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone="MBA8MA197"
/clone_1lb="Ascaris suum (parasitic nematode) body wall
muscle and hypodermis"
/sex="mixed"
/tissue.type="body wall muscle and hypodermis"
/dev_stage="adult"
/note="Vector: lambda zap II; Site_1: NotI (5'end);
Site_2: NotI (3'end); Ascaris suum is an intestinal
nematode parasite of pigs. The library was constructed
from dissected body wall muscle and hypodermis tissue for
Dr. T. Geary, Pharmacla-Upjohn Inc, Kalamazoo, MI, USA
[tgeary@em.pnu.com]"
BASE COUNT 144 a 205 c 179 g 115 t 1 others
ORIGIN

Query Match 2.1%; Score 39.2; DB 9; Length 644;
Best Local Similarity 51.1%; Pred. No. 5.9;
Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtctctctgagcactctcccccagagagtcgtgtgacagagcagccagccta 1107
DB 228 ATGACCTCCGGGAGACACAGCAAGATGCTGCCGGGAGAGACCTGCCGCCAG 287

QY 1108 gcaatctgagcagcaacagtcctctctgacacctctccagagagccactgtgtc 1167
DB 288 GACCTCCAGGACGAGATGCCAAGCTTCACGATCTCTGCGCCGTACACACAGATGTC 347

QY 1168 agagtcagagatgcctggaacatccgagtcctgagcagcagcagagctagtgc 1227
DB 348 CCTGTCAAGCTCCTCTGAGACTCTGCGCTCTGTGACACCTGAGAGATGACACAC 407

RESULT 13
BI282815/c 549 bp mRNA linear EST 19-JUL-2001
LOCUS
DEFINITION
UT-R-DA0-byp-e-09-0-UI-R-DA0 Rattus norvegicus cDNA clone
BI282815
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
1 (bases 1 to 549)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized rat eye library cDNA library preparation: M.B.
Soares Lab Clone distribution: cDNA clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.
Location/Qualifiers
1. 549

FEATURES

source

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DA0-byp-e-09-0-UI"
/clone_1lb="UI-R-DA0"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DA0
library is a non-normalized library constructed from rat
salivary gland tissue. For a detailed description of the
library from which this clone was derived, please visit
our web site at ratat.eng.uiowa.edu. The subtraction has
been previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_L1B-UI-R-DA0
TAG_TISSUE-rat eye
TAG_SEQ-CAGCC"
BASE COUNT 55 a 118 c 201 g 175 t
ORIGIN

Query Match 2.1%; Score 39; DB 10; Length 549;
Best Local Similarity 46.2%; Pred. No. 6.4;
Matches 129; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 882 ggaacagcaacagatctctctgacactctcccccagagcactgtgtcagggagac 941
DB 496 GGCCACACGACGAGACACCTCAGGGCCACCCACACGAGAGACCAAGAGACTCA 437

QY 942 gccatgctagcaatcccgagcagcaacagatctctctgacactctcccccaggggtc 1001
DB 436 GGCCACCCCCACCAAGAGAGACACAGAGACTCTGAGGCCACCCACCAAGAGGC 377

QY 1002 acttggtcagagcagcagcgcctgacactcctgagcagcaacagtcctctctgc 1061
DB 376 CCACAGCAGAGACCTCAAGGCCCAACCCACAGAGAGACCAAGAGACTCAAGGC 317

QY 1062 actcttccccagagtcactgtgtcagagcagcagcagcctgacactcctggagac 1121
DB 316 CCACCCCCACCAAGAGAGACCAAGAGACTCTGAGGCCCAACCCACCAAGAGACCA 257

QY 1122 caacagtcctctctgacactctctccagagcact 1160
DB 256 CAGCAGAGACTCAAGGCCCAACCCCAAGAGTGGCCT 218

RESULT 14
CNS020KA 699 bp DNA linear GSS 14-MAY-2000
LOCUS
DEFINITION
Tetradon nigroviridis genome survey sequence PUC-ori end of clone
153D20 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION
AL206803
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS: genome survey sequence.
Tetradon nigroviridis.
Tetradon nigroviridis

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2002, 00:45:01 ; Search time 3374.16 Seconds
(without alignments)
11733.843 Million cell updates/sec

Title: US-09-755-456-1
Perfect score: 1830
Sequence: 1 gaattcagatgcctcatacc.....tcctctagaagtgtgggac 1830

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 1081749327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US0954_COMB.seq.*

```

Result NO.	Score	Query Match	Length	ID	Description
1	1830	100.0	1830	US-09-755-456-1	Sequence 1, Appl1
2	176	9.6	1116	US-09-755-456-3	Sequence 3, Appl1
3	52.8	2.9	2000	US-09-887-272A-5283	Sequence 5283, Ap
4	50.8	2.8	2000	US-09-887-272A-5283	Sequence 5283, Ap
5	45.4	2.5	740	PCT-US00-14864A-26	Sequence 26, Appl
6	45.4	2.5	740	US-09-950-083-995	Sequence 95, App
7	42.6	2.3	533	US-09-181-996-561	Sequence 561, App
8	42.2	2.3	1011	US-09-360-207-36034	Sequence 36034, A
9	42.2	2.3	1564	US-09-950-083-7697	Sequence 7697, Ap
10	41.4	2.3	516	US-09-196-718-9735	Sequence 675, App
11	41.4	2.3	516	US-09-196-718-9735	Sequence 3736, Ap
12	40.8	2.2	1084	US-09-610-154-72	Sequence 72, Appl
13	40.8	2.2	1084	US-09-938-301-72	Sequence 14, Appl
14	40.6	2.2	7218	US-08-466-194-14	Sequence 270, App
15	40	2.2	470	US-09-162-47-8274	Sequence 2574, Ap
16	40	2.2	470	US-09-162-47-8274	Sequence 2813, Ap
17	40	2.2	832	US-09-621-976-2813	Sequence 2813, Ap
18	40	2.2	832	US-09-621-976-2813	Sequence 1380, Ap
19	40	2.2	2791	US-09-360-207-1380	Sequence 3, Appl1
20	39.8	2.2	4064	US-09-818-656-3	Sequence 1150, Ap
21	39.8	2.2	660	US-09-188-162-1150	Sequence 26, Appl
22	39.2	2.1	708	US-10-001-873-26	Sequence 521, App
23	39.2	2.1	1026	US-09-605-700-521	Sequence 101912, Sequence 101912,
24	38.8	2.1	467	US-09-528-409-101912	Sequence 101912,
25	38.6	2.1	436	US-09-933-524-101912	Sequence 3986, Ap
26	38.6	2.1	436	US-09-933-524-101912	Sequence 13854, A
27	38.6	2.1	436	US-09-721-589-3986	Sequence 17426, A
28	38.6	2.1	587	PCT-US01-00663-13854	Sequence 8047, Ap
29	38.2	2.1	611	US-09-864-761-17426	
30	38.2	2.1	611	US-09-864-761-17426	
31	38.2	2.1	611	US-09-864-761-17426	
32	38.2	2.1	611	US-09-864-761-17426	
33	38.2	2.1	611	US-09-864-761-17426	
34	38.2	2.1	611	US-09-864-761-17426	
35	38.2	2.1	611	US-09-864-761-17426	
36	38.2	2.1	611	US-09-864-761-17426	
37	38.2	2.1	611	US-09-864-761-17426	
38	38.2	2.1	611	US-09-864-761-17426	
39	38.2	2.1	611	US-09-864-761-17426	
40	38.2	2.1	611	US-09-864-761-17426	
41	38.2	2.1	611	US-09-864-761-17426	
42	38.2	2.1	611	US-09-864-761-17426	
43	38.2	2.1	611	US-09-864-761-17426	
44	38.2	2.1	611	US-09-864-761-17426	
45	38.2	2.1	611	US-09-864-761-17426	
46	38.2	2.1	611	US-09-864-761-17426	
47	38.2	2.1	611	US-09-864-761-17426	
48	38.2	2.1	611	US-09-864-761-17426	
49	38.2	2.1	611	US-09-864-761-17426	
50	38.2	2.1	611	US-09-864-761-17426	
51	38.2	2.1	611	US-09-864-761-17426	
52	38.2	2.1	611	US-09-864-761-17426	
53	38.2	2.1	611	US-09-864-761-17426	
54	38.2	2.1	611	US-09-864-761-17426	
55	38.2	2.1	611	US-09-864-761-17426	
56	38.2	2.1	611	US-09-864-761-17426	
57	38.2	2.1	611	US-09-864-761-17426	
58	38.2	2.1	611	US-09-864-761-17426	
59	38.2	2.1	611	US-09-864-761-17426	
60	38.2	2.1	611	US-09-864-761-17426	
61	38.2	2.1	611	US-09-864-761-17426	
62	38.2	2.1	611	US-09-864-761-17426	
63	38.2	2.1	611	US-09-864-761-17426	
64	38.2	2.1	611	US-09-864-761-17426	
65	38.2	2.1	611	US-09-864-761-17426	
66	38.2	2.1	611	US-09-864-761-17426	
67	38.2	2.1	611	US-09-864-761-17426	
68	38.2	2.1	611	US-09-864-761-17426	
69	38.2	2.1	611	US-09-864-761-17426	
70	38.2	2.1	611	US-09-864-761-17426	
71	38.2	2.1	611	US-09-864-761-17426	
72	38.2	2.1	611	US-09-864-761-17426	
73	38.2	2.1	611	US-09-864-761-17426	
74	38.2	2.1	611	US-09-864-761-17426	
75	38.2	2.1	611	US-09-864-761-17426	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

32	38	2.1	1501	75	US-60-360-207-17623	Sequence 17623, A
C 33	38	2.1	1501	31	US-09-815-264-50372	Sequence 50372, A
C 34	38	2.1	15955	24	US-09-620-392-58240	Sequence 58240, A
C 35	38	2.1	5044	24	US-09-620-392-65702	Sequence 65702, A
C 36	38	2.1	8419	24	US-09-620-392-99885	Sequence 99885, A
C 37	38	2.1	10829	24	US-09-620-392-30253	Sequence 30253, A
C 38	38	2.1	85882	24	US-09-620-392-48150	Sequence 48150, A
C 39	38	2.1	94641	28	US-09-702-134-7928	Sequence 7928, A
C 40	38	2.1	94641	31	US-09-815-264-79401	Sequence 79401, A
C 41	37.8	2.1	678	24	US-09-634-3068-121331	Sequence 121331, A
C 42	37.8	2.1	771	24	US-09-634-3068-150013	Sequence 150013, A
C 43	37.8	2.1	2727	33	US-09-887-2728-3090	Sequence 3090, A
C 44	37.8	2.1	2727	71	US-60-325-448-694	Sequence 694, A
C 45	37.6	2.1	411	16	US-09-293-972-31298	Sequence 31298, A

ALIGNMENTS

```

RESULT 1
US-09-755-456-1
: Sequence 1, Application US/09755456
: GENERAL INFORMATION:
: APPLICANT: DELBAC, FREDERIC
: APPLICANT: DANCHIN, ANTOINE
: APPLICANT: VIVARES, CHRISTIAN
: TITLE OF INVENTION: MICROSOPORDIIN POLAR TUBE PROTEINS, NUCLEIC ACIDS
: TITLE OF INVENTION: CODING FOR THESE PROTEINS AND THEIR APPLICATIONS
: FILE REFERENCE: 1566-00
: CURRENT APPLICATION NUMBER: US/09/755,456
: CURRENT FILING DATE: 2001-01-05
: PRIOR APPLICATION NUMBER: PCT/FR99/01630
: PRIOR FILING DATE: 1999-07-06
: PRIOR APPLICATION NUMBER: FR 98/08692
: PRIOR FILING DATE: 1998-07-07
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 1
: LENGTH: 1830
: TYPE: DNA
: ORGANISM: Encephallitozoon cuniculi
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (345)..(1529)
: US-09-755-456-1

```

Query Match	100.0%	Score 1830	DB 29	Length 1830
Best Local Similarity	100.0%	Pred. NO. 0		
Matches 1830	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	gaattcagatgcctcatcatttggaattaaaatctgattgtcatctgttatatacctg	60		
Db 1	gaattcagatgcctcatcatttggaattaaaatctgattgtcatctgttatatacctg	60		
QY 61	ggcggaacggcgcgctgattcttcgaagggtgtgtgcctaccagattgcacgaagttcc	120		
Db 61	ggcggaacggcgcgctgattcttcgaagggtgtgtgcctaccagattgcacgaagttcc	120		
QY 121	ggaagtgctcttgatgataagaaagtaagccattgtgtgtcttcatacatgtcatcgccct	180		
Db 121	ggaagtgctcttgatgataagaaagtaagccattgtgtgtcttcatacatgtcatcgccct	180		
QY 181	ttcggtcgtttccccaagaatccaattatctctccaggaacttcacacctcagaatgaaa	240		
Db 181	ttcggtcgtttccccaagaatccaattatctctccaggaacttcacacctcagaatgaaa	240		
QY 241	caggagatcaaacctctctgtgcaaatcgtatagatatcgattggaagaatcgtaaacacggag	300		
Db 241	caggagatcaaacctctctgtgcaaatcgtatagatatcgattggaagaatcgtaaacacggag	300		
QY 301	tttgaataaaaaagataataacctccgaaaaacgacagatttaagatgaaaggtattctcta	360		
Db 301	tttgaataaaaaagataataacctccgaaaaacgacagatttaagatgaaaggtattctcta	360		

Db	301	ttcgaaataaagatataaataactcccgaaaaacgacagatttaagatgaaggtatttcta	360
Qy	361	agatccctctctgctctatctatctgcctgatagtatggaagaatgltatltacgaacacgcac	420
Db	361	agatccctctctgctctatctatctgcctgatagtatggaagaatgltatltacgaacacgcac	420
Qy	421	tgttcgacgaatgcataatgtgcttaactccgvgacacaaggtatgtgtcagacgcgtcgt	480
Db	421	tgttcgacgaatgcataatgtgcttaactccgvgacacaaggtatgtgtcagacgcgtcgt	480
Qy	481	atgtctgtatccccaagcacccgggaaacaaatagcaaaactgtgcaagcgtttcaacggaa	540
Db	481	atgtctgtatccccaagcacccgggaaacaaatagcaaaactgtgcaagcgtttcaacggaa	540
Qy	541	caatatctctctctcccgctgacccaacaaatctccagttgaatcccgvgaaataatagcga	600
Db	541	caatatctctctctcccgctgacccaacaaatctccagttgaatcccgvgaaataatagcga	600
Qy	601	atgagacatctccatccgctctctcgcgaagaatgtatggaacaaatgcagaatgtccgtatga	660
Db	601	atgagacatctccatccgctctctcgcgaagaatgtatggaacaaatgcagaatgtccgtatga	660
Qy	661	agcaacttcgacgacacacaggaataaataatcaaggagacacacacaggtccatgtgaa	720
Db	661	agcaacttcgacgacacacaggaataaataatcaaggagacacacacaggtccatgtgaa	720
Qy	721	ccccagagacgaacacagccttctgtcagtatctccacacacctctgcgtacccgtgacgt	780
Db	721	ccccagagacgaacacagccttctgtcagtatctccacacacctctgcgtacccgtgacgt	780
Qy	781	tggagctctgcacagctcccatctctgtgtgcagatgttctcgtctgtctacccaacag	840
Db	781	tggagctctgcacagctcccatctctgtgtgcagatgttctcgtctgtctacccaacag	840
Qy	841	ccgttcacagcgtctactacaacaaatggaataatcccgatattctctgacaaacacagatcttt	900
Db	841	ccgttcacagcgtctactacaacaaatggaataatcccgatattctctgacaaacacagatcttt	900
Qy	901	ctggcacctctccccaagagacacttctgtgcagaggaacagccatgcttaacatctcgt	960
Db	901	ctggcacctctccccaagagacacttctgtgcagaggaacagccatgcttaacatctcgt	960
Qy	961	gacacgaacagaatccttctctgacatctctccccaaggggtcacttgtgtcagggagacag	1020
Db	961	gacacgaacagaatccttctctgacatctctccccaaggggtcacttgtgtcagggagacag	1020
Qy	1021	ccaagcctatgacactctctgvgacgaacacaggtccttctcgtgcacacttccccaagatca	1080
Db	1021	ccaagcctatgacactctctgvgacgaacacaggtccttctcgtgcacacttccccaagatca	1080
Qy	1081	ctttgtgtcagaaggaacagccacgcctctagacatctcctgggacagaacaaagatctcttctgca	1140
Db	1081	ctttgtgtcagaaggaacagccacgcctctagacatctcctgggacagaacaaagatctcttctgca	1140
Qy	1141	ccctctccccaagagacacttctgtgtcagatcaaaagtatgctctggaacatcccgagttc	1200
Db	1141	ccctctccccaagagacacttctgtgtcagatcaaaagtatgctctggaacatcccgagttc	1200
Qy	1201	ctggacacgacaggaacagttctatgtgacaggtgttctgcccctccaaagattccaaacccgttca	1260
Db	1201	ctggacacgacaggaacagttctatgtgacaggtgttctgcccctccaaagattccaaacccgttca	1260
Qy	1261	tgcgcacatccatgaaactatgtggaatagggtatcccttctctcaacgataacgcccac	1320
Db	1261	tgcgcacatccatgaaactatgtggaatagggtatcccttctctcaacgataacgcccac	1320
Qy	1321	acctggaatcaatggaatctctgtgtgtgacatacagaagaacgvggggacatccctcgac	1380
Db	1321	acctggaatcaatggaatctctgtgtgtgacatacagaagaacgvggggacatccctcgac	1380
Qy	1381	aaaaacccgagagatccgcacacgagtalgcacatgagagcctgtgcacaacacaccaa	1440
Db	1381	aaaaacccgagagatccgcacacgagtalgcacatgagagcctgtgcacaacacaccaa	1440

Qy	1441	cggtatattataggaacaacgagtagtactctgttgcgcacgaagtgtacaaatgcaatlaact	1500
Db	1441	cggtattatattaggcaacaacgagtagtactctgttgcgcacgaagtgtacaaatgcaatlaact	1500
Qy	1501	ctccagtcgaacacagtcgcttccaaatgtctgttgcgcataaaataaagagtttaacttctt	1560
Db	1501	ctccagtcgaacacagtcgcttccaaatgtctgttgcgcataaaataaagagagtttaacttctt	1560
Qy	1561	tctcttcgctcttcttggaagagcttggaatggagtgagagtcctacggcgctgaagtgaat	1620
Db	1561	tctcttcgctcttcttggaagagcttggaatggagtgagagtcctacggcgctgaagtgaat	1620
Qy	1621	gccacaacctctctctgcgcgaagaacacatctcgagtgcttctctctgttgcgcagagtggtg	1680
Db	1621	gccacaacctctctctgcgcgaagaacacatctcgagtgcttctctctgttgcgcagagtggtg	1680
Qy	1681	taaaagagattcccgagagatttagcagcttggagtagacatgattgaatcagtatataac	1740
Db	1681	taaaagagattcccgagagatttagcagcttggagtagacatgattgaatcagtatataac	1740
Qy	1741	tctcccaaatatttattactctctctgatttatatcccgagccaatctggaagaatgcttc	1800
Db	1741	tctcccaaatatttattactctctctgatttatatcccgagccaatctggaagaatgcttc	1800
Qy	1801	gaattcaagctcccttagaagtgtaggac	1830
Db	1801	gaattcaagctcccttagaagtgtaggac	1830

RESULT 2
US-09-755-456-3
; Sequence 3, Application US/09755456

```

1  APPLICANT: DELBAC, FREDERIC
2  APPLICANT: DANCHIN, ANTOINE
3  APPLICANT: VIARES, CHRISTIAN
4  TITLE OF INVENTION: MICROSPORDIAN POLAR TUBE PROTEINS, NUCLEIC ACIDS
5  TITLE OF INVENTION: CODING FOR THESE PROTEINS AND THEIR APPLICATIONS
6  FILE REFERENCE: 1566-00
7  CURRENT APPLICATION NUMBER: US/09/755,456
8  CURRENT FILING DATE: 2001-01-05
9  PRIOR APPLICATION NUMBER: PCT/FR99/01630
10 PRIOR FILING DATE: 1999-07-06
11 PRIOR APPLICATION NUMBER: FR 98/08692
12 PRIOR FILING DATE: 1998-07-07
13 NUMBER OF SEQ ID NOS: 10
14 SOFTWARE: PatentIn Ver. 2.1
15 SEQ ID NO 3
16 LENGTH: 1116
17 TYPE: DNA
18 ORGANISM: Encephalitozoon intestinalis
19 FEATURE:
20 NAME/KEY: CDS
21 LOCATION: (1)..(1113)
22 US-09-755-456-3

```

Query Match	9.68;	Score 176;	DB 29;	Length 1116;
Best Local Similarity	53.58;	Pred. No. 6.6e-42;		
Matches 635;	Conservative 0;	Mismatches 460;	Indels 93;	Gaps 9

Oy	345	atgaaaggattcttcaaaatccctcttgcgcctatttcccgatgaagttgaaatgctc	400
Db	1	atgaaaggattcttctaaagctctctctcagccctatttgcctcaatggaagtgtgaagtgctc	60
Oy	405	tatttcagcaaacccgcacgtgtgcagcaatgacatatagtcctaactccggagacaacaggatgtg	460
Db	61	tatttcacaactgtgctgtgtgtgcagatctc-----aaacacaaagactctcaaggaga--	100
Oy	465	gtctcagcgcgcgtcgtatgtgtctatccccgcacccccgggaaacatatagcacaactgtgca	520
Db	110	-caaccacacgtctatatgtctgtgtctctcgtgtgcacacagaaacatatggccaactgtgaa	160

QY	525	agcggtctcacaggaacacatactctccctcccgctgcacacacatctccagtgactccg	584
Db	169	tacatgcctcaagaacacatgatactgcctctctactactactacatgactccacagtgcca	228
QY	585	gggaaaactagcgagaaatgagatact-----cactgcgtccctgcaggaagatgta	635
Db	229	ggcacaactcgtgagaaacggagaaacactactctccacaactctctccagagaaatgtg	288
QY	636	ggaacatgcacaagattgcgcgatctgaaagcactcgcagacccacggaaacaaatccagggacg	695
Db	289	ggaaacatgcacaagattgctgtgttgtaaagcatctgtgatagcaccggaaacatcatcaaacact	348
QY	696	aacacaaaggtctcaagggcctctgtgaaaccccaagagcagaaacagcctctgtcagtgatctcc	755
Db	349	tgcgaacccggaaacagaaacttgcgcctcact-----cagccagtagcagctacaatt	399
QY	756	aacacatcctgcgcgaacccggagactgtggaggtctcgcacagatctccatctgtgtgcagat	815
Db	400	ggccacacaactggtgtgtgtctctctgcagaaacgcgcgaagcagctgttacaactcttact	459
QY	816	gttcctctgtctgcacacccacccagcagatctccagcgtactacaacaaatgaaacatccggt	875
Db	460	ccaaagcgcgctctctgcgcgcacggcggaacaocat---attctccatctaaacccggaccc	515
QY	876	atctctggagacagcaacagatctctctctctgcgaactctctcccccaggaagccaattgtgtcag	935
Db	516	agggtactaataatgtgtgcaattccccgggcgaataacttcaacagtgaaatgtctctccacag	575
QY	936	ggacagagccatgtccctagcaactctctgtgacagcaacagatctctctctgcgaactctcccca	995
Db	576	tgcctctctcttgcacaagtggtgcgcgcggaaacaaagaaagctccaccccacagcagctacc	635
QY	996	gggggtcaacttgtgtcagggagcagcgcacgcgctcagcactctctgtggcagcaaacaggtctct	1055
Db	636	aggcgcgtgtctcatctcgtgaaacccaactctctgtgcacaatagtaacagggaaactcaag-----	689
QY	1056	tcgtgcacactctcccccaggaagtcactcttgtgtcagggagacagccagcgcctagcaactct	1115
Db	690	-ttagcggaacaaaccccctgtgacagcaactctctgcgcggagaaatcgtctccgtgtggaaagctcc	747
QY	1116	gggcagagcaaacaggtctctctctgtgcacacccctctccagagacacacttgtgtcagagatcaa	1175
Db	748	cagcgcgcatcaagtaacttctgtgagacccttaacccc-----	782
QY	1176	ggtatgccttggaaacatccggaggtctcctgtgcagcagcagggacagltcagtgtgaaacgtgtgt	1235
Db	783	-----ttctctgttagccaagaacaaactctgtgacagca-----atgcgtctgc	822
QY	1236	ggccctcagtttccaacacccctgtgcacggcgcaatccatcaataatagtggaatgtgagat	1295
Db	823	actctccccaatacaaaaacccctgtcaatgtgcacactccctaaatggtatcagaacgtaaatgtgagat	882
QY	1296	cctctcttcaacgcatalcagccccaacacctctcgcagatctggaatctctgtgtgtgacatacag	1355
Db	883	cccagctccaactgtgtagcggcccaaccccttggacaatcttgagcccttgatctgcagacag	942
QY	1356	aagacgggggggagacatctctgcgcagcaaaaaacccggaaagttccgcgaacgagtatgcacatg	1415
Db	943	aagctc-----aacatcatctctgcgcagcaaaaaaagaaacccctgtgacagataltgtgaatg	996
QY	1416	gagcgctgtgcacaacccaacaaacaggttatatatagtgaaacagcaggaataltctgttga	1475
Db	997	gaagcatagcgctgcacaacactccaactgtcttctcaggaatagcgtcgagatctctccttagc	1056
QY	1476	ccaggaatgtacaatgtcaatlaactctccatctgcagacaactgtctgcacaa	1523
Db	1057	ccggaggtgtataattcactcaactcactcatcacaagaacgtctgtgcacaa	1104

RESULT 3
US-09-887-272A-5263
; Sequence 5263, Application US/09887272A

```

/ GENERAL INFORMATION:
/ APPLICANT: Hou, Yu-Ming
/ APPLICANT: Quan, Sheng
/ APPLICANT: Chang, Hur-Song
/ APPLICANT: Zhu, Tong
/ APPLICANT: Whitham, Steve
/ APPLICANT: Golf, Steve
/ APPLICANT: Glazebrook, Jane
/ APPLICANT: Chen, Menglong
/ APPLICANT: Katagiri, Fumiaki
/ APPLICANT: Xie, Zhiyi
/ APPLICANT: Tao, Yi
/ APPLICANT: Zou, Guangzhou
/ APPLICANT: Cooper, Bret
/ TITLE OF INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST
/ FILE REFERENCE: 1360.003052
/ CURRENT APPLICATION NUMBER: US/09/887,272A
/ CURRENT FILING DATE: 2001-06-23
/ PRIOR APPLICATION NUMBER: 60/213,634
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/214,926
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/261,320
/ PRIOR FILING DATE: 2001-01-12
/ PRIOR APPLICATION NUMBER: 60/264,353
/ PRIOR FILING DATE: 2001-01-26
/ PRIOR APPLICATION NUMBER: 60/273,879
/ PRIOR FILING DATE: 2001-03-07
/ PRIOR APPLICATION NUMBER: 09/887,271
/ PRIOR FILING DATE: 2001-06-22
/ NUMBER OF SEQ ID NOS: 6813
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5263
/ LENGTH: 2000
/ TYPE: DNA
/ ORGANISM: Oryza sativa
US-09-887-272A-5263

Query Match      2.9%; Score 52.8; DB 33; Length 2000;
Best Local Similarity 9.1%; Pred. NO. 0.00046;
Matches 93; Conservative 457; Mismatches 464; Indels 6; Gaps 3;

Oy 773 ggtgactgtgagctgtgcacagctcatctgtgtgcagctgtgtcgtgtgtccta 832
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 7 gwksrctswgrirmymagmmscarngssrsmksmgskymkrcscgkcmktrrkskv 66
Oy 833 ccacccagcagctccagcgtactacaacaatggaacatccggtatctccttgagacaaca 892
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 67 ysasasagrtgskwssgysgkqmkrrykrskirwgrirgrrmrrsmrwmgryrcaarsg 126
Oy 893 gatccttctgagcacttcccccagagcacttctgttcagggagacagcctagctag 952
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 127 rmagsggrmmgskrmsymwcyarogcckrkkskysgwgkccrrgaagsgagsgakyl 186
Oy 953 cactcctgagacacacagatccttctgtgacacttcccccagggagccttggtgta 1012
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 187 ksgsmakrmmsacgrsgcrrasrygtrskgtlykmtlyssarcmraymttsysw 246
Oy 1013 gggagcagggcagcctagcactcctcttgagcagcaagcgttcttcttgacacttccccc 1072
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 247 acssytwcstkrismnmkrmrkmrwsrsgwyswskmmmtctaykksysrwy---my 302
Oy 1073 aggaagcacttctgtcagagcagcagcctagcactccctcgsgagcaagaagctcct 1132
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 303 tgggwtgatltywrgysmrnammykkmwyrygkmgmkrgwagrrmmrsmcrrskacyym 362
Oy 1133 ttctggacacctctccagagcagcacttctgtcagagatcagagatagcttgacacatc 1192
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 363 twmwmtrrrrrwakkssrtrrkkhkwcmrkkykrmgysmrmsckrarrwmkrcrsgr 422
Oy 1193 cggagttccttgagcagcagcagcagctcagtgagcagctgtgtgcctccatccaa 1252

```

```

Db 423 awkmgcrgcmctcrmksgymmrwkswkrmaskkykmaarmyrrwkkkcosrtlmwgtktrgmm 482
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 1253 ccctgtcatgcccgcacatcatgacaatagtggaatggtgtcctctctc-accgact 1311
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 483 glmgicrrykrksgmkrrckrrrrgrrymrmtwkrmysatylmrcarkkysaarkarc 542
Oy 1312 acagcccaaccctcgatcactggaactcctgtgtgtacatacagaagcggggagacat 1371
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 543 wrygkygywagmmkrykkmymykrkyskcsyckmsyysacmskarskaganc 602
Oy 1372 cctgagcagaacaaaccggaagctccgccaagcagatagccatgagagccttgcaaac 1431
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 603 kskmsawskmsrckckcaskrsakrryamgmgtsgmsrskstcywkrwksmk 662
Oy 1432 caaccacaaggtattatagcaacagcagagatcttcttgacccaagaatgtaaatg 1491
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 663 stctmymyskylayagygyryrwcymymrwyryryrsymlymawysltlmaatg 722
Oy 1492 caattaactctcatgcaacactgctgtccaatgctgctagcctaataaacaagatlt 1551
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 723 mkysgrywtsvlykycsw-kyrsmlyyswwaktkmwrlyatrmmyrysmkyt 781
Oy 1552 aatcttcttctcctcgctccttcttgaaagctgagatggagagagctatggctg 1611
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 782 wclmgywmywrrlynmkmykckctlywsatlygttwaawmaktkmrmgmtgakt 841
Oy 1612 aatgaaatgcacaactctctctgccaagaacatcgatgctctcctctgagca 1671
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 842 rgarararwwkwalwcalwcalrwmckgkgaawaakwrryyswmawyyktrrt 901
Oy 1672 ggaagttgtlaacagagatcccgagatlagcagccttgagagtaacatgaatla 1731
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 902 ryltcmwkarwsgwayrwmwkskmmwmmvkggrgwtkywytcttkmccgatrakym 961
Oy 1732 gtattaaactctcaaatatttattcttctgtttatattccagagcaactgagaa 1791
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 962 ccagwamyswltcrltywmltrwmlwasarlakrarmmklrawskysaraywkmaga 1021

RESULT 4
US-09-887-272A-5263/c
/ Sequence 5263, Application US/09887272A
/ GENERAL INFORMATION:
/ APPLICANT: Hou, Yu-Ming
/ APPLICANT: Quan, Sheng
/ APPLICANT: Chang, Hur-Song
/ APPLICANT: Zhu, Tong
/ APPLICANT: Whitham, Steve
/ APPLICANT: Golf, Steve
/ APPLICANT: Glazebrook, Jane
/ APPLICANT: Chen, Menglong
/ APPLICANT: Katagiri, Fumiaki
/ APPLICANT: Xie, Zhiyi
/ APPLICANT: Tao, Yi
/ APPLICANT: Zou, Guangzhou
/ APPLICANT: Cooper, Bret
/ TITLE OF INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST
/ FILE REFERENCE: 1360.003052
/ CURRENT APPLICATION NUMBER: US/09/887,272A
/ CURRENT FILING DATE: 2001-06-23
/ PRIOR APPLICATION NUMBER: 60/213,634
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/214,926
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/261,320
/ PRIOR FILING DATE: 2001-01-12
/ PRIOR APPLICATION NUMBER: 60/264,353
/ PRIOR FILING DATE: 2001-01-26
/ PRIOR APPLICATION NUMBER: 60/273,879
/ PRIOR FILING DATE: 2001-03-07
/ PRIOR APPLICATION NUMBER: 09/887,271

```

```

: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 6813
: SOFTWARE: FASTSEQ for Windows Version 4.0.
: SEQ ID NO: 5263
: LENGTH: 2000
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-887-272A-5263

```

Query Match 2.8%; Score 50.8; DB 33; Length 2000;
Best Local Similarity 9.4%; Pred. No. 0.0019;
Matches 89; Conservative 424; Mismatches 426; Indels 7; Gaps 3

QY	207	atctcccaagacttcaaccctcagaatgaagaagaatgaactctgttgaac	266
Db	958	ATYCGTMMMAAGRRMRMAMCWCMMKKMMKMTSCMMKKYMTWTSWCTTMMGCAIRYAY	899
QY	267	gtagatctcgatltvgagacatltgaaccagagaglttgaataaagataatactcc	326
Db	988	AMRRRRRTYKWSWNRBMWTMTKKAAMTMMTCMCAKKNYAATGATMMMYTYTYCYAMT	839
QY	327	gaaacgcagag-cttaagatgaaggtattcttaagatccctctgtcctatggcct	385
Db	838	CACKKYKMAATKMTTWTACAWRATSNRWRMAGMRKKRTKMKRAYMRWRMCKMKGARW	779
QY	386	gatgaatltvgagaatgtctattcaagaaacgcagctgttgaacaaatgatgtgctaac	445
Db	778	MKSRYRMKKKKYATRYKWMAMTMMWMSNRMSYRHMGMGRMMSAVRYSRBAKCAKT	719
QY	446	tcgggagcaacaggtatgtctcagcagcgctgtatgtctgatalccccaagaccggg	505
Db	718	KYASASAMTKRAKSYTYRRRMWTKKKGWTYRRYSRCSRMTPARRSKRRKMAAGSMSC	659
QY	506	aaccatgacaactgttgcagcggttccacagagacatatctctctcccgctgaac	565
Db	658	WMYRRGARSMWYSKYSCSACCKKTTYMTMSVMS-----TGYTGMYSYKSMSTMYSKY	604
QY	566	cacatctccagtgactccggggaacacatgacgaagaatgagacatctccatcgctctgc	625
Db	603	MGMKTCMTMYSMKSTRRSMGMSQMSRMYMMWMMKKMRKRYMRMKCTWRRCMYR	544
QY	626	agaagatgtgaagaacatgcagaattccglatgaagacatctgagacgcagcaaac	685
Db	543	WGYTMTTSSNRMMYTRRYARTSKRRMYTYRRKYCTTYTGMTMCKSMRIGCYKA	484
QY	686	atcagggagacacacaggtcagcgcttgttgaaccacagagcagcaacagctttgc	745
Db	483	CKKCCYAMCWMKAAYSGMAMTYRRKYSMMHMSTKYMSMWYKCKSMKYGAKGTCCKMW	424
QY	746	agtgatctccaacactctctcgcttaccggtgaactgtggaatgtcagtcagcttcac	805
Db	423	TYCGYGMKWTYMGSKYSRCYKYMWMYTKGMWMTMYTSAYSSMMTYTYTYAYTYMWY	364
QY	806	tgctcaggttgcctctgtctgtcacaacacagcagatltcagagctactacaacatg	865
Db	363	YKRRTGTSWGSKSKKRYCTWCMCYMCMCRIMRRKMRKKRTKYSKRCYCMRATYTMCC	304
QY	866	aacatccggtattctctvgagacaaacagatccttctvgaaccttccccagag-cca	924
Db	303	YRKRGWYSRRSMNTACKMWRMSMRSCBSYSMTKYMKMMKKSYYMSGARBSGTWSR	244
QY	925	cttggttcaaggagcagcatgctcagcactcttgaacagaacagatcctttgtga	984
Db	243	SAAKRTYKGYSTRAKMMAACMYISACRYSTNTSYCCGCTCGSSKMYMKSCSMRWT	184
QY	985	ctcttcccccaagggtacttgttgaagaaagcagcagctagaacatcttggcagc	1044
Db	183	CSMSCCSYTCYGCAMCMSCMSMWTYMGSCGYTGMRKMRKYSMCKKYGCSCTKYSY	124
QY	1045	aacaggtctcttcgggaactcttcccccaagagactcttgttcaaggagcagcagc	1104
Db	123	TGYRYTCMWTYSYKYCYCYCYWYMSYMMYMMKCMCSRSCSSMMSACVSTSSRMS	64

```

Oy 1105 ctgacactcctgycagcaacagtccttctgcaaccctctccc 1150
      :: : | :: : : : : : | : : : :
Db 63 MYAAKMGMCSSGMYRMSKSCMKMYSKYSSCAKXTGSKKCTKRKY 18

```

```

RESULT 5
PCT-US00-14964A-26
: Sequence 26, Application PC/TUS0014964A
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: 48 Human Secreted Prob
: FILE REFERENCE: P5360PCT
: CURRENT APPLICATION NUMBER: PCT/US00/14964A
: CURRENT FILING DATE: 2000-06-01
: PRIOR APPLICATION NUMBER: 60/138,627
: PRIOR FILING DATE: 1999-06-11
: NUMBER OF SEQ ID NOS: 170
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 26
: LENGTH: 740
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US00-14964A-26

```

Query Match	2.58;	Score 45.4;	DB 1;	Length 740;
Best Local Similarity	47.48;	Pred. No. 0.048;		
Matches 136;	Conservative 0;	Mismatches 151;	Indels 0;	Gaps 0;

Oy	898	tttttggaactctccccccagagagcaacttgtgtacaggagaaaggccagccttagcctac	957
Db	349	tttggggaactctccccccagagcccaagccttggtctcgaagtcgtctcttcacagcg	408
Oy	958	cttgacagcaaacagatcctcttcgtgcactctcccccagaggtcactttgtgtcagagac	1017
Db	409	tggccacatbatgctctgctccgcgtcactctctctgcttagtgcacaggagatctgcttc	468
Oy	1018	aggcgacagccttagagacacccctggcgagcaaaaggtctctctctgcaactctcccccagag	1077
Db	469	aggacagctccctctgttaagcttcaacaatgatactagctcaagtcacatctctctaaccaatg	528
Oy	1078	tcaatttgtgtcagaagagacagcagcagctgaactctctctgtggacgaacacaggtcttctc	1137
Db	529	gcaacatgtgtctcatcatcagagacatcatcgatctgtgtgcacaacatgagtgctagtcag	588
Oy	1138	gcaaccttctcccaagagagcaacttgtgtacagataaagatgatgct	1184
Db	589	tcaagctctccagcagcagcttagcaatgtgctctctctctctctctctctctctctctct	635

```

1  RESULT 6
2  US-09-950-083-995
3  ; Sequence 995, Application US/09950083
4  ; GENERAL INFORMATION:
5  ; APPLICANT: Rosen, et. al
6  ; TITLE OF INVENTION: Human Secreted Proteins
7  ; FILE REFERENCE: PS805
8  ; CURRENT APPLICATION NUMBER: US/09/950, 083
9  ; CURRENT FILING DATE: 2001-09-12
10 ; PRIOR APPLICATION NUMBER: 60/278, 650
11 ; PRIOR FILING DATE: 2001-03-27
12 ; PRIOR APPLICATION NUMBER: 09/833, 245
13 ; PRIOR FILING DATE: 2001-04-12
14 ; PRIOR APPLICATION NUMBER: PCT/US01/11988
15 ; PRIOR FILING DATE: 2001-04-12
16 ; PRIOR APPLICATION NUMBER: PCT/US00/06043
17 ; PRIOR FILING DATE: 2000-03-09
18 ; PRIOR APPLICATION NUMBER: PCT/US00/06012
19 ; PRIOR FILING DATE: 2000-03-09
20 ; PRIOR APPLICATION NUMBER: PCT/US00/06058
21 ; PRIOR FILING DATE: 2000-03-09
22 ; PRIOR APPLICATION NUMBER: PCT/US00/06044

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
222

QY 1122 caacag 1128
Db 172 GCACAG 166

RESULT 9

US-09-950-083-7697
Sequence 7697, Application US/09950083
GENERAL INFORMATION:
APPLICANT: Rosen, et. al
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS805
CURRENT APPLICATION NUMBER: US/09/950,083
PRIOR FILING DATE: 2001-09-12
PRIOR FILING DATE: 2001-03-27
PRIOR FILING DATE: 09/833,245
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US01/11988
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US00/06043
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06012
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06058
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06044
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06059
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06042
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06014
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06013
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06049
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06057
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06824
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06765
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06792
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06830
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06782
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06822
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06791
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06781
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/07505
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07440
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07506
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07507
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07535
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07525
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07534

PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07483
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07526
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07527
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07661
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07579
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07723
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07724
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/14929
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/07722
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07578
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07726
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07677
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07725
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/09070
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08982
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08983
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09067
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09066
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09068
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08981
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08980
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09071
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09069
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/15136
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14926
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14963
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/15135
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14934
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14933
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/15137
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14928
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14973
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14964
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/26376
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US00/26371
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US00/26324
PRIOR FILING DATE: 2000-09-26

PRIOR APPLICATION NUMBER: PCT/US00/26323
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US00/26337
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US01/13318
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/124,146
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/167,061
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/124,093
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/166,989
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/124,145
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/168,654
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/124,099
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/168,661
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/124,096
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/168,622
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/124,143
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/168,663
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/124,095
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/138,598
PRIOR FILING DATE: 1999-06-11
PRIOR APPLICATION NUMBER: US 60/168,665
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/125,360
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: US 60/138,626
PRIOR FILING DATE: 1999-06-11
PRIOR APPLICATION NUMBER: US 60/168,662
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/124,144
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/138,574
PRIOR FILING DATE: 1999-06-11
PRIOR APPLICATION NUMBER: US 60/168,667
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/124,142
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/138,597
PRIOR FILING DATE: 1999-06-11
PRIOR APPLICATION NUMBER: US 60/168,666
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/125,359
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: US 60/168,664
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/126,051
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/169,906

Query Match 2.38; Score 42.2; DB 35; Length 1564;
Best Local Similarity 46.78; Pred. No. 0.7; Mismatches 153; Indels 0; Gaps 0;

Db 898 ttcttgacactctcccccagagacactttgtlcaaggagacatgcttagcactc 957
|| |||| ||||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 509 ttctggcagcttcacagcagccatgagctgtgctcaggggtgtcttcacgscg 568
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 958 ctgagcagcaacagatccttctgacactctcccccaggggttaactttgtcaggagac 1017
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 569 tggccacataagctgtgctcggctcactcttctgcttaggtaccagagatctcgtc 628

Qy 1018 aggcacagccttagcactcttggcagcaacaggtccttcttgcgactcttcccccagag 1077
||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 629 aggacagctcctctgtctacagctacatgatatctcagctcagttacttaccatag 688
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 1078 tcaacttgtlcaaggagacagcagccttagcactcttggcagcaacaggtcttctg 1137
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 689 gcaacatgttctacatcaagacagctacatgagctgttgccaacatgagcttagcag 748
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 1138 gcaacttctccagagagcactttgttcaagatcaaggtatgct 1184
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 749 tcaacttccagcagcagctgcttgccttgcctcaagtaactt 795
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

RESULT 10
US-60-196-718-675
Sequence 675, Application US/60196718
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000456
CURRENT APPLICATION NUMBER: US/60/196,718
CURRENT FILING DATE: 2000-04-13
NUMBER OF SEQ ID NOS: 7494
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 675
LENGTH: 516
TYPE: DNA
ORGANISM: HUMAN
US-60-196-718-675

Query Match 2.38; Score 41.4; DB 58; Length 516;
Best Local Similarity 45.78; Pred. No. 0.66;
Matches 144; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 863 tggacacacggtattctctgagcagcaacagatccttcttgcactcttcccccagagc 922
||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 121 tggcgtgcctcctcactgctgtacacatgagcagagccttgtgtcttccacagctgt 180
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 923 caatttgtlcaaggagacagcagccttagcactcttgaagacagacatccttctg 982
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 181 gcatatgagcagagagccttgcctgcctcactctgttgcacatgagtgctggagcctg 240
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 983 cacttctcccccaggggtcactttgtlcaaggagacagcagccttagcactcttggca 1042
||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 241 cactgcccctcactgctgtgacagctgagcttgagccttgcctcactcactgtgca 300
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 1043 gcaacaggtccttcttgcgactcttcccccaggggtacactttgttcaaggagacagc 1102
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 301 cgtgagctgagcagcctgtgcctcactgctgtgacagctgagcgtggagccttgcgc 360
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 1103 gcttagcactcttggcagcagaggtccttcttgcagcccttctccagagacacttt 1162
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 361 tgcctcactgtgtgacagcgtgagcgttgagccttgcctcactcactgtgtgacgt 420
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 1163 gtlcagatcaagg 1177
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 421 ggaagctgagccttg 435
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

RESULT 11
US-60-196-718-3736
Sequence 3736, Application US/60196718
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000456
CURRENT APPLICATION NUMBER: US/60/196,718

```
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3736
; LENGTH: 516
; TYPE: DNA
; ORGANISM: HUMAN
US-60-196-718-3736
```

```
Query Match          2.3%; Score 41.4; DB 58; Length 516;
Best Local Similarity 45.7%; Pred. No. 0.66;
Matches 144; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
```

```
QY 863 tggacatccggtattctcgtgagacaagatccttctgacacttcccccagagc 922
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 tggcgtccctcactcgtgtacacatgagcagcgctggtctctactctgt 180
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 923 cacttgtgtcagagcagcagcagcagcagcagcagcagcagcagcagcagc 982
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 gcatatgagcagagcagcagcagcagcagcagcagcagcagcagcagcagc 240
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 983 cacttccccccaggggtacacttctgtcagagcagcagcagcagcagcagcagc 1042
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 cactgcccctcactcgtgtcagcagcagcagcagcagcagcagcagcagcagc 300
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1043 gcaacaggtcttctcgtgacacttcccccagagcagcagcagcagcagcagc 1102
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 cgtgagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1103 gcttagcactcctcgtgagcagcagcagcagcagcagcagcagcagcagc 1162
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 361 tgcctcactcgtgtcagcagcagcagcagcagcagcagcagcagcagcagc 420
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1163 gtgtcagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1177
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 421 ggaagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 435
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 12
US-09-610-154-72
; Sequence 72, Application US/09610154
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-155
; CURRENT APPLICATION NUMBER: US/09/610,154
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-610-154-72
```

```
Query Match          2.2%; Score 40.8; DB 23; Length 1084;
Best Local Similarity 49.5%; Pred. No. 1.5;
Matches 105; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```

```
QY 905 cactcttccccagagcagcagcagcagcagcagcagcagcagcagcagcagc 964
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 562 cagcttcagcgccgagcagcagcagcagcagcagcagcagcagcagcagcagc 621
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 965 gcaacagatccttctcgtgacacttcccccaggggtcagcagcagcagcagc 1024
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 622 cagcgtctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagc 681
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1025 gcttagcactcctcgtgagcagcagcagcagcagcagcagcagcagcagc 1084
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 682 tcttctggttttagtaacaacagcttagctcagggcagctcctctgtttacagtcagcat 741
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 1085 gtgtcagagcagcagcagcagcagcagcagcagcagcagcagcagc 1116
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 742 ggaatcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 773
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 13
US-09-938-301-72
; Sequence 72, Application US/09938301
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-155
; CURRENT APPLICATION NUMBER: US/09/938,301
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 09/610,154
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-301-72
```

```
Query Match          2.2%; Score 40.8; DB 35; Length 1084;
Best Local Similarity 49.5%; Pred. No. 1.5;
Matches 105; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```

```
QY 905 cactcttccccagagcagcagcagcagcagcagcagcagcagcagcagcagc 964
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 562 cagcttcagcgccgagcagcagcagcagcagcagcagcagcagcagcagcagc 621
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 965 gcaacagatccttctcgtgacacttcccccaggggtcagcagcagcagcagc 1024
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 622 cagcgtctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagc 681
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1025 gcttagcactcctcgtgagcagcagcagcagcagcagcagcagcagcagc 1084
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 682 tcttctggttttagtaacaacagcttagctcagggcagctcctctgtttacagtcagcat 741
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1085 gtgtcagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1116
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 742 ggaatcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 773
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 14
US-08-466-194-14
; Sequence 14, Application US/08466194
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, NW, Suite 500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; APPLICATION NUMBER: US/08/466,194
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/232,463
; FILING DATE: 22-APR-1994
```



```

; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9848
; LENGTH: 1564
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-105-299-9848

```

Query Match 2.3%; Score 42.2; DB 7; Length 1564;

Best Local Similarity 46.7%; Pred. No. 0.072; Mismatches 153; Indels 0; Gaps 0;

Matches 134; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```

Qy 898 ttcttgcaactcttccccagaggaagcactttgttcaaggagacgcatgcttagcactc 957
    || |||| |||| || |||| || |||| || |||| || |||| || |||| || ||||
Db 509 ttcttgagcagctctccagctacagcccatatggtctgctcaagggtgtcttccacgcg 568
Qy 958 ctggagacgaacagatccttcttgcaactcttccccaggggtcactttgttcaaggagc 1017
    || || || || || || || || || || || || || || || || || || || || ||
Db 569 tggcccatatgagctcgtgcctcgttcaactcttcttgcttaggtcacagagagctcgcctc 628
Qy 1018 aggcacagcttagaactcctctgggagcaacaggtccttcttgcaactcttccccagagag 1077
    || || || || || || || || || || || || || || || || || || || || ||
Db 629 aggcacagctcctctgcttcaaggtcacaatgatactagctacagttcctcttccacatgg 688
Qy 1078 tcaactttgttcaaggagcagcgccctagcaactccttggcagcaacaggtccttcttg 1137
    || || || || || || || || || || || || || || || || || || || || ||
Db 689 gcaacatggtcttaccatcagagacatgacagctgtggccaacatgagctagctcagag 748
Qy 1138 gaacacttctccagagagcactttgttcaagatcaagtgatgcct 1184
    || || || || || || || || || || || || || || || || || || || || ||
Db 749 tcaacttccagcagcagctacagatgcttctgttcaagtcagctt 795

```

RESULT 3

```

; US-09-818-656A-74/C
; Sequence 74, Application US/09818656A
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001191
; CURRENT APPLICATION NUMBER: US/09/818,656A
; CURRENT FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (301)..(301)
; OTHER INFORMATION: Either A or C may be present, or both may be
; OTHER INFORMATION: absent (single base deletion)
; US-09-818-656A-74

```

Query Match 2.2%; Score 40.8; DB 5; Length 601;

Best Local Similarity 45.5%; Pred. No. 0.12; Mismatches 168; Indels 0; Gaps 0;

Matches 141; Conservative 1; Mismatches 168; Indels 0; Gaps 0;

```

Qy 889 aacagatccttcttgcaactcttccccagaggaagcactttgttcaaggagacagcatgc 948
    || || || || || || || || || || || || || || || || || || || || ||
Db 322 ACCAGTGTCCACCCCTGTTCCTTCCACACGATGGGGTGTGTGGATTCATCCACACCATTC 263

```

```

Qy 949 ctgacactcttgagacagcaacagatccttcttgcaactcttccccaggggtcactttgt 1008
    || || || || || || || || || || || || || || || || || || || || ||
Db 262 CCTTAGCCACATGCTCTTACAGCTCCAGCTGTTCCCTTCCAGCCAGTGCCT 203
Qy 1009 gtcagaggaagcagcagccttagcactcttgagcagcaacagatccttcttgcaactctc 1068
    || || || || || || || || || || || || || || || || || || || || ||
Db 202 ACCAGCTCCACCCCTGTTCCTTCCACACATGAGGTGTGTGGATTCATCCACACCATTC 143
Qy 1069 ccccaagagtaactttgttcaaggagcagcagcagccttagcactcttggcagcaacag 1128
    || || || || || || || || || || || || || || || || || || || || ||
Db 142 CCTTAGCCACCCGCTCTTACAGCTCCAGCTGTTCCCTTCCAGCTGAGGTGTGGA 83
Qy 1129 tcttcttgcaactcttccccagaggaagcactttgttcaagatcaagtgatgcctggaa 1188
    || || || || || || || || || || || || || || || || || || || || ||
Db 82 TCCATTCCTCCACACATTCCTTCCAGCCAGCCGCTTACAGCTCCACCCGTTCCCTGGA 23
Qy 1189 catcggaggt 1198
    || || || || || || || || || || || || || || || || || || || || ||
Db 22 CACGTGGGGT 13

```

RESULT 4

```

; US-10-123-155-10/C
; Sequence 10, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvairet, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Daniel
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-123-155-10

```

Query Match 2.2%; Score 40.6; DB 7; Length 594;

Best Local Similarity 6.9%; Pred. No. 0.14; Mismatches 266; Indels 0; Gaps 0;

Matches 35; Conservative 183; Mismatches 266; Indels 0; Gaps 0;

```

Qy 859 acaatggaacatcggatctccttgagcagcaacagatccttcttgcaactcttccccag 918
    || || || || || || || || || || || || || || || || || || || || ||
Db 565 A.MCTT..AMM.M.CBT..STT.YA.M.YT.S.S.SYSYSYS.S.S.SDSYSYA.SY 506
Qy 919 gagcaactttgttcaaggagcagcagcagccttagcactcttggaaagcaacagatcctt 978
    || || || || || || || || || || || || || || || || || || || || ||
Db 505 STS.S.SMSYSYSSSDDY.CYCYRYRHCSDSYSYSYV..CRCCYVT..SYSRYCHVSC 446
Qy 979 ctggcagcttccccaggggtcactttgttcaaggagcagcagcagccttagcactctt 1038
    || || || || || || || || || || || || || || || || || || || || ||
Db 445 CCSDYCYYSRYSRYSYSYSYSYSYTDYCSRRCCYSYSYSYSYSYSYSYSYSATST 386
Qy 1039 ggcagcaacaggtccttcttgcaactcttccccagagatcactttgttcaaggagcag 1098

```

```

Db 385 SSSSSSSSSYYTSTNC.T.CC...T..MCAABCSTTTTTTTTTT..HSCC.SA..A. 326
OY 1099 ccaagcctagcactcctcggcagcaagagtcctcttcggcaccctctcccaaggagca 1158
Db 325 M..YC.A.SYSYS.S.S.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.Y.YSY 266
OY 1159 cttgtgctagcagatcagatccttcggaacatccggagtcctcggacagagagcagt 1218
Db 265 YCSRKT...TWTDM.T.T..MHY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
OY 1219 ctgtgacagctgtgtgccccctcagattccaaccctgtcatcgcgcacatgaaca 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146
OY 1279 ttagtgaagatggtatcctctctcctacagcagcccaaccctcggactcggat 1338
Db 145 .CCR..H.R..R.G...S.CT.HN.B.CYR.RNGMY.HS...S.AA...ARH..DH. 86
OY 1339 cctgtgtgacatacagaagcgg 1362
Db 85 DTYBH.YH.KNNAHC.W.KKNB.S 62

```

```

RESULT 5
US-10-137-871-10/c
; Sequence 10, Application US/10137871
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-871-10

```

```

Query Match 2.28; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.98; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;
OY 859 acaatggaacatccggtatcctcgtgacagcaagatccttcggcactctcccccag 918
Db 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSIA.SY 506
OY 919 gaggcacttgtgtcagagcagcagcatgctcctagcactcctgacagcaagatcctt 978
Db 505 SYS.S.S.SMSYSYSSSDDY.CYCCYRHHCSYSYSYY.CRCCTYT.SYSRYDCHYSC 446
OY 979 ctggcactctcccccgggggactcctgtgtcagggagcagcagcctcgtcactcgt 1038

```

```

Db 445 CCSDYCYCYSYRYSYSYSYSYSYDYSYRCCCYVYSSSSSSSSAYST 386
OY 1039 ggcagcaacaggtcctcttcgtgacactctcccccagagtcacttgtgtcaggagcag 1098
Db 385 SSSSSSSSSYYTSTNC.T.CC...T..MCAABCSTTTTTTTTTT..HSCC.SA..A. 326
OY 1099 ccaagcctagcactcctcggcagcaagagtcctcttcggcaccctctcccaaggagca 1158
Db 325 M..YC.A.SYSYS.S.S.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.Y.YSY 266
OY 1159 cttgtgctagcagatcagatccttcggaacatccggagtcctcggacagagagcagt 1218
Db 265 YCSRKT...TWTDM.T.T..MHY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
OY 1219 ctgtgacagctgtgtgccccctcagattccaaccctgtcatcgcgcacatgaaca 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146
OY 1279 ttagtgaagatggtatcctctcctacagcagcccaaccctcggactcggat 1338
Db 145 .CCR..H.R..R.G...S.CT.HN.B.CYR.RNGMY.HS...S.AA...ARH..DH. 86
OY 1339 cctgtgtgacatacagaagcgg 1362
Db 85 DTYBH.YH.KNNAHC.W.KKNB.S 62

```

```

RESULT 6
US-10-141-761-10/c
; Sequence 10, Application US/10141761
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-10

```

```

Query Match 2.28; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.98; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;
OY 859 acaatggaacatccggtatcctcgtgacagcaagatccttcggcactctcccccag 918
Db 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSIA.SY 506
OY 919 gaggcacttgtgtcagagcagcagcatgctcctagcactcctgacagcaagatcctt 978
Db 505 SYS.S.S.SMSYSYSSSDDY.CYCCYRHHCSYSYSYY.CRCCTYT.SYSRYDCHYSC 446

```

OY	979	ctgacgactctcccccaggaggtactctgtgtcaggaagacgacacgctgacactctg	1038
Db	445	CCSYYCYTSTSYSTNYSYSTSWSTSYTYDCTSRKCCCTYSYSTSSYSTSSAST	386
OY	1039	gacgacacagctctctctctgacactctcccccaggatcattgtgtcagggacag	1098
Db	385	SSSSSSSSSYTSTNMC.T.CC...T.MCAACSTTTTTTTTTT.HSCC.SA.A.	326
OY	1099	ccacgcttagcactcctctggcagcaaacagctcctctctgacaccctctccagagcca	1158
Db	325	M.YC.A.SYSYS.SSS.S.SYMR.HRA.SHYTTRS..S.MCY.YM.Y..Y.YYSY	266
OY	1159	cttgtgtcaggtcaagaatgacgcctggagacatccggagctctctgacagcagggacgt	1218
Db	265	YCSAKTM...TMDM.T.T..MHMY.KTB.HCHTKCPAT.MN.HTB...N..HBHB.B	206
OY	1219	ctaattgacagctgtgtgcccctcagaattccaaacctgatacgcgcacatcatalgaaca	1278
Db	205	H.H...B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y.	146
OY	1279	ttagtgaatagtgatctctctctcaccgacatacagcccaaacctcgatcattggat	1338
Db	145	.CCR..H.R..R.G...S.CT..HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH.	86
OY	1339	ccgtgtgtgacatacagaagaagc	1362
Db	85	DTYBH..YH.KNNAH.C.W.KNB..S.62	

```

RESULT 7
US-10-140-864-10/c
: Sequence 10. Application US/10140864
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerltsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Collin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P9330R1C184
: CURRENT APPLICATION NUMBER: US/10/140,864
: CURRENT FILING DATE: 2002-05-07
: Prior Application removed - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 10
: LENGTH: 594
: TYPE: PR1
: ORGANISM: Homo Sapien
: US-10-140-864-10

```

Query Match	2.28;	Score 40.6;	DB 7;	Length 594;
Best Local Similarity	6.98;	Pred. No. 0.14;		
Matches 35;	Conservative 183;	Mismatches 286;	Indels 0;	Gaps 0;

Oy 859 acatagtacaatccgtattctccttgacagcaacagatccttctgycactcttccccag 918
| : | : | : : : : : : : : :
Db 565 A.MCTT..AMM.M.CBT.STT.YA.W.YT.S.S.S.SVSYSYS.S.S.SDSYSTA.SY 506

Oy 919 gaccactctgtgtcgaaggacagccatcctcctagacactctccttgagacagaacagatcctt 978
 Db 505 SYS.S.SWSSTSYSSSDY.CCCTCYRHCSDBSYSTYY.CRCCTYT.SISRUDCHASC 446
 Oy 979 ctggcactctctcccccagggatctactctgtgtcgaaggacagccacgcctagactccty 1038
 Db 445 CCSDYCCSYSTSYRXYSTYSYSSMSYSTYDCCSYRCCCYXSYSSSSSYSSASVT 386
 Oy 1039 ggcgcgaacaggtctcttcctcctgcgcactctcccccaggaatctactctgtgtcgaaggacag 1098
 Db 385 SSSSSSSSSSYSTNYTNC.T.CC...T.MCABCSSTTTTTTTTTT..HSCC.SA.A. 326
 Oy 1099 ccaagcctagacactcctcgtggcagaacaagttcctctctgcacactctcccaagagcca 1158
 Db 325 M..YC.A.SYSSYS.SSS.S.SYMR.HRA.SHYTIRS..S.MYCY.YM.Y..Y.YYSY 266
 Oy 1159 cttgtgtctcagatcaagaatgatcgtcctggaacatccgagatcttcctggaacagcaggacagt 1218
 Db 265 YCSRKTM...TMTDM.T.T..MHMY.KYB.HCHTKCRAT.MN.HTB...N..HBHB.B 206
 Oy 1219 ctatgtgacagtggtgtgccccctcagatcccaaacctctatcgcgcacatcatgacaa 1278
 Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASB.RARC.STMM.T.MMM.H.S.Y. 146
 Oy 1279 ttatgtgaaatgggtatcctctctctctacagccatcagcccaaacctcggatcatcggagt 1338
 Db 145 .CCR..H.R..R.G...S.CT.HN.B.CVRY.RNGMY.HS...S.AA...ARH..DH. 86
 Oy 1339 cctgtgttgacatcagaagaagc 1362
 Db 85 DTYBH.YH.KNNAH.C.W.KNB.S.62

```

RESULT      8
US-10-140-923-10/c
; Sequence 10, Application US/10140923
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Deenoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140,923
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-923-10

```

Query match 2.28; score 40.6; DB 7; length 594;

Best Local Similarity: 0.36; Freq: NO: 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

QY 859 acaatggaacatccggtattccttgacagcaacagatccttctgcactcttcccccag 918

Db 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SY 506
QY gagccacttctgtcagagacagccatcgtcagacacccctcgcagacagacagacagcctt 978
Db 505 SYS.S.SMSYSYSSSDDY.CYCCYRHCSDSYSTYY.CRCCTYY.SYSRDCCHSC 446
QY ctgcagactctcccccaggggacattgtgtcagagagacagccagcctcagcactctg 1038
Db 445 CCSDYYCYSTRYSRYSYSYSYSYTDYCSYRCCYSSSSSSSSSAVST 386
QY 1039 ggcagcaacaggtcctctcagcactctcccccagagacattgtgtcagagagcag 1098
Db 385 SSSSSSSSYTSTNYC.T.CC...T.MCABCSSTTTTNTT.T.HSCC.SA.A. 326
QY 1099 ccagcctcagacactcctcgtcagacagagcctcctcgcagacccctccagagagca 1158
Db 325 M..YC.A.SYSYS.SSS.S.SYMR.HRA.SHYTRS.S.MCY.YM.Y.YY.YSY 266
QY 1159 ctgtgtcagagatcaggtatgctcgtgaacatccgagctcgtcgtcagacagagcagt 1218
Db 265 YCSRKT...TMTDM.T.T..MHY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
QY 1219 ctgtgtcagagatcaggtatgctcgtgaacatccgagctcgtcgtcagacagagcagt 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MM.H.S.Y. 146
QY 1279 ttatggaatcaggtatcctctcctcagacacagcccaacccgtcagcagcagat 1338
Db 145 .CCR..H.R..R.G...S.CT.HN.B.CYR.RNMY.HS...S.AA...ARRH.DH. 86
QY 1339 cctgtgtcagacacagagcag 1362
Db 85 DTYBH.YH.KNNAMC.W.KKNB.S 62

RESULT 9
US-10-141-756-10/c
; Sequence 10, Application US/10141756
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C200
; CURRENT APPLICATION NUMBER: US/10/141.756
; PRIOR APPLICATION DATE: 2002-05-08
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-756-10

Query Match 2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. No. 0.14;

Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;
QY 859 acaatgagacatccggtatctcgtgacagcagacacatcctcctcgtcagcactcctcccccag 918
Db 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SY 506
QY gagccacttctgtcagagacagccatcgtcagacacccctcgcagacagacagacagcctt 978
Db 505 SYS.S.SMSYSYSSSDDY.CYCCYRHCSDSYSTYY.CRCCTYY.SYSRDCCHSC 446
QY ctgcagactctcccccaggggacattgtgtcagagagacagccagcctcagcactctg 1038
Db 445 CCSDYYCYSTRYSRYSYSYSYSYTDYCSYRCCYSSSSSSSSSAVST 386
QY 1039 ggcagcaacaggtcctctcagcactctcccccagagacattgtgtcagagagcag 1098
Db 385 SSSSSSSSYTSTNYC.T.CC...T.MCABCSSTTTTNTT.T.HSCC.SA.A. 326
QY 1099 ccagcctcagacactcctcgtcagacagagcctcctcgcagacccctccagagagca 1158
Db 325 M..YC.A.SYSYS.SSS.S.SYMR.HRA.SHYTRS.S.MCY.YM.Y.YY.YSY 266
QY 1159 ctgtgtcagagatcaggtatgctcgtgaacatccgagctcgtcgtcagacagagcagt 1218
Db 265 YCSRKT...TMTDM.T.T..MHY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
QY 1219 ctgtgtcagagatcaggtatgctcgtgaacatccgagctcgtcgtcagacagagcagt 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MM.H.S.Y. 146
QY 1279 ttatggaatcaggtatcctctcctcagacacagcccaacccgtcagcagcagat 1338
Db 145 .CCR..H.R..R.G...S.CT.HN.B.CYR.RNMY.HS...S.AA...ARRH.DH. 86
QY 1339 cctgtgtcagacacagagcag 1362
Db 85 DTYBH.YH.KNNAMC.W.KKNB.S 62

RESULT 10
US-10-141-759-10/c
; Sequence 10, Application US/10141759
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C197
; CURRENT APPLICATION NUMBER: US/10/141.759
; PRIOR APPLICATION DATE: 2002-05-08
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-759-10


```

; SEQ ID NO 10
; LENGTH: 594
; TYPE: PR
; ORGANISM: Homo Sapien
US-10-140-805-10

```

```

Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

```

```

OY 859 acaatgagacatccggtatctcttgagacagcaacatctcttgacactctcccccag 918
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 565 A.MC.TT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SY 506

OY 919 gagccacttgytcaaggacagccatgcttagacactccggcagacagacagactctt 978
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 505 SYS.S.SMSYSYSSSDDY.CYCCYRHCSDSYSYSY.Y.CRCCT.YT.SYSRYDCHISC 446

OY 979 ctggacactctcccccagggtcacttgytgcaggagacagccagcctagcactctg 1038
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 445 CCSDYCYYSYSRYSYSYSYSYSYSYTDYCSYRCCCYSSYSSSSSAYST 386

OY 1039 ggcagacagtgctctcttgacactctcccccaggagtcacttgytgcaggagcag 1098
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 385 SSSSSSSSYTSTNYC.T.CC...T.MCAABCSITTTTTTTTTT..HSCC.SA..A. 326

OY 1099 ccagcctagacactccggcagacagagtcctcttgacacccctctccagagacga 1158
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 325 M.YC.A.SYSYS.S.SS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y..Y.YYSY 266

OY 1159 ctgtgctcagatcaaggtatgctctggaacatccggaggtcctgagacagagcag 1218
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 265 YCSKRTM...TMTDM.T.T..MHMY.KYB.HCHTKCRAT.MN.HTB...N..HBHB.B 206

OY 1219 ctatgagacagtgtgtgcccctcagatccaaacccgtcactgcccgcacatcaga 1278
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STTM.T.MMM.H.S.Y. 146

OY 1279 ttagtgaatggtatctctcttctacgcatacagcccaaacctcgatcagtgat 1338
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 145 .CCR..H.R.R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH. 86

OY 1339 cctgtgtgacatacagaagcgg 1362
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 85 DTYBH.YH.KNNAH.C.W.KKNB.S 62

```

```

RESULT 13
US-10-142-885-10/c
; Sequence 10, Application US/10142885
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C248
; CURRENT APPLICATION NUMBER: US/10/142,885

```

```

; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PR
; ORGANISM: Homo Sapien
US-10-142-885-10

```

```

Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

```

```

OY 859 acaatgagacatccggtatctcttgagacagcaacatctcttgacactctcccccag 918
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 565 A.MC.TT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SY 506

OY 919 gagccacttgytcaaggacagccatgcttagacactccggcagacagacagactctt 978
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 505 SYS.S.SMSYSYSSSDDY.CYCCYRHCSDSYSYSY.Y.CRCCT.YT.SYSRYDCHISC 446

OY 979 ctggacactctcccccagggtcacttgytgcaggagacagccagcctagcactctg 1038
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 445 CCSDYCYYSYSRYSYSYSYSYSYSYTDYCSYRCCCYSSYSSSSSAYST 386

OY 1039 ggcagacagtgctctcttgacactctcccccaggagtcacttgytgcaggagcag 1098
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 385 SSSSSSSSYTSTNYC.T.CC...T.MCAABCSITTTTTTTTTT..HSCC.SA..A. 326

OY 1099 ccagcctagacactccggcagacagagtcctcttgacacccctctccagagacga 1158
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 325 M.YC.A.SYSYS.S.SS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y..Y.YYSY 266

OY 1159 ctgtgctcagatcaaggtatgctctggaacatccggaggtcctgagacagagcag 1218
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 265 YCSKRTM...TMTDM.T.T..MHMY.KYB.HCHTKCRAT.MN.HTB...N..HBHB.B 206

OY 1219 ctatgagacagtgtgtgcccctcagatccaaacccgtcactgcccgcacatcaga 1278
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STTM.T.MMM.H.S.Y. 146

OY 1279 ttagtgaatggtatctctcttctacgcatacagcccaaacctcgatcagtgat 1338
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 145 .CCR..H.R.R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH. 86

OY 1339 cctgtgtgacatacagaagcgg 1362
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 85 DTYBH.YH.KNNAH.C.W.KKNB.S 62

```

```

RESULT 14
US-10-146-731-10/c
; Sequence 10, Application US/10146731
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C248
; CURRENT APPLICATION NUMBER: US/10/142,885

```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-10

```

```

Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. NO. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

```

```

QY 859 acaatgaaacatccggtatcctctgagcaagaacagatccttcttgacactcttccccag 918
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSTSYS.S.S.SDSYSTA.SY 506

QY 919 gagcacattgtgcaaggagccatgctacgactcctctgagcaagaacatcctt 978
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 505 SYS.S.S.SMSYSYSSSDY.CYCCYRHCSDSYSYTY.CRCCTYT.SYRHDCHYSC 446

QY 979 ctggacactctccccagggtacttctgtcaggagcaagccagcctagacactctg 1038
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 445 CCSDYCYSYSTRYSYSYSYSYSYSYTDYCSYRCCCYSSSYSSSYSSSYSSAYST 386

QY 1039 ggcagcaacagtccttcttgacactctccccagagatcattgtgtcaggagcag 1098
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 385 SSSSSSSSYTYSTNYC.T.CC...T.MCAABCSSTTTTTTTTTT..HSCC.SA..A. 326

QY 1099 ccagcctagacactcctctgagcaagaagtccttcttgacactcttccccagagca 1158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 325 M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.YYSY 266

QY 1159 ctgtgtcaggaatgaagtaagctctgagcaacatccgagtcctctgagcagaggaag 1218
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 265 YCSRKT...TWTDM.T.T..MHMT.KYB.HCHTRCRAAT..NN..HTB...N..HBHB.B 206

QY 1219 ctagtgcagagtgctgtgccccctcagatcccaaacctgtcagccgcacatccatga 1278
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146

QY 1279 ttagtgaatggatccttctctaccgcacatacagcccaaacctcgatcagtgat 1338
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 145 .CCR..H.R..R.G...S.CT..HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH. 86

QY 1339 cctgtgtgacatacagaagcag 1362
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 85 DTYBH.YH.KNNAHC.W.KKNB.S.62

```

```

RESULT 15
US-10-142-426-10/c
; Sequence 10, Application US/10142426
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K

```

```

; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C224
; CURRENT APPLICATION NUMBER: US/10/142,426
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-426-10

```

```

Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. NO. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

```

```

QY 859 acaatgaaacatccggtatcctctgagcaagaacagatccttcttgacactcttccccag 918
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSTSYS.S.S.SDSYSTA.SY 506

QY 919 gagcacattgtgcaaggagccatgctacgactcctctgagcaagaacatcctt 978
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 505 SYS.S.S.SMSYSYSSSDY.CYCCYRHCSDSYSYTY.CRCCTYT.SYRHDCHYSC 446

QY 979 ctggacactctccccagggtacttctgtcaggagcaagccagcctagacactctg 1038
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 445 CCSDYCYSYSTRYSYSYSYSYSYSYTDYCSYRCCCYSSSYSSSYSSSYSSAYST 386

QY 1039 ggcagcaacagtccttcttgacactctccccagagatcattgtgtcaggagcag 1098
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 385 SSSSSSSSYTYSTNYC.T.CC...T.MCAABCSSTTTTTTTTTT..HSCC.SA..A. 326

QY 1099 ccagcctagacactcctctgagcaagaagtccttcttgacactcttccccagagca 1158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 325 M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.YYSY 266

QY 1159 ctgtgtcaggaatgaagtaagctctgagcaacatccgagtcctctgagcagaggaag 1218
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 265 YCSRKT...TWTDM.T.T..MHMT.KYB.HCHTRCRAAT..NN..HTB...N..HBHB.B 206

QY 1219 ctagtgcagagtgctgtgccccctcagatcccaaacctgtcagccgcacatccatga 1278
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146

QY 1279 ttagtgaatggatccttctctaccgcacatacagcccaaacctcgatcagtgat 1338
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 145 .CCR..H.R..R.G...S.CT..HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH. 86

QY 1339 cctgtgtgacatacagaagcag 1362
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 85 DTYBH.YH.KNNAHC.W.KKNB.S.62

```

```

Search completed: September 1, 2002, 02:10:33
Job time: 6037 sec

```